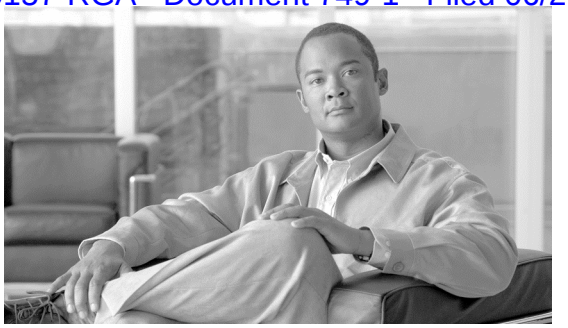
Nourbakhsh Dec. Exhibit 1 (Part 1 of 2)

Cisco Unified Contact Center Enterprise Solution Reference Network Design (SRND)

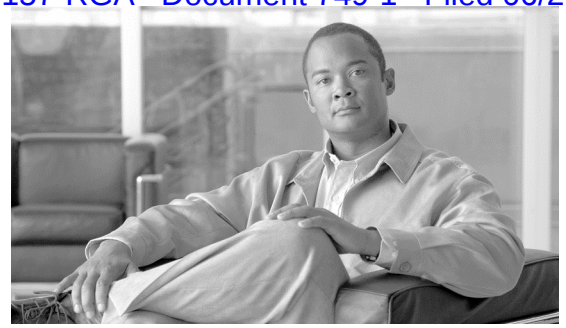
Cisco Unified Contact Center Enterprise (Unified CCE) Release 8.0
June 26, 2012

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

Cisco Unified Contact Center Enterprise Solution Reference Network Design



CONTENTS



Preface

This document provides design considerations and guidelines for deploying Cisco Unified Contact Center Enterprise Release 8.0 and later releases in a Cisco Unified Communications System.

This document builds upon ideas and concepts presented in the latest version of the *Cisco Unified Communications SRND Based on Cisco Unified Communications Manager*, which is available online at:

http://www.cisco.com/en/US/products/sw/voicesw/ps556/products_implementation_design_guides_list.html

This document assumes that you are already familiar with basic contact center terms and concepts and with the information presented in the *Cisco Unified Communications SRND*. To review IP Telephony terms and concepts, refer to the documentation at the preceding URL.

New or Changed Information for This Release

Note Unless stated otherwise, the information in this document applies to Cisco Unified Contact Center Enterprise Release 8.0.

The following chapters contain information that has changed significantly from previous releases of Cisco Unified Contact Center Enterprise.

Revision History

This document may be updated at any time without notice. You can obtain the latest version of this document online at:

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_implementation_design_guides_list.html

Visit this Cisco.com website periodically and check for documentation updates by comparing the revision date on the front title page of your copy with the revision date of the online document.

The following table lists the revision history for this document.

Revision Date	Comments
May 26, 2011	Updated content for RSM to Agent Phone (RTP).
May 17, 2011	Corrected some minor errors.
May 24, 2010	Corrected several minor errors.
March 19, 2010	Initial version of this document for Cisco Unified Contact Center Enterprise Release 8.0.
November 30, 2009	Corrected some minor errors.
August 18, 2009	Content was updated.
May 7, 2009	Minor update for a change in Computer Telephony Integration (CTI) capacity limits.
April 22, 2009	Content was updated for Cisco Unified Communications System Release 7.1.
November 12, 2008	Corrected several minor errors.
October 29, 2008	Revised some of the sizing information for Cisco Unified Contact Center Enterprise components and servers, and added a chapter on Cisco Unified Contact Center Management Portal (Unified CCMP).
August 27, 2008	Initial version of this document for Cisco Unified Contact Center Enterprise Release 7.5.

Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, submitting a service request, and gathering additional information, see the monthly *What's New in Cisco Product Documentation*, which also lists all new and revised Cisco technical documentation, at:

http://www.cisco.com/en/US/docs/general/whatsnew/whatsnew.html

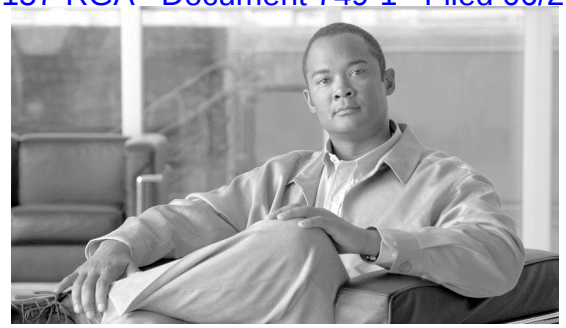
Subscribe to the *What's New in Cisco Product Documentation* as a Really Simple Syndication (RSS) feed and set content to be delivered directly to your desktop using a reader application. The RSS feeds are a free service and Cisco currently supports RSS version 2.0.

Cisco Product Security Overview

This product contains cryptographic features and is subject to United States and local country laws governing import, export, transfer and use. Delivery of Cisco cryptographic products does not imply third-party authority to import, export, distribute, or use encryption. Importers, exporters, distributors and users are responsible for compliance with U.S. and local country laws. By using this product you agree to comply with applicable laws and regulations. If you are unable to comply with U.S. and local laws, return this product immediately.

Further information regarding U.S. export regulations may be found at:

http://www.access.gpo.gov/bis/ear/ear_data.html



CHAPTER 1

Architecture Overview

Cisco Unified Contact Center Enterprise (Unified CCE) is a solution that delivers intelligent call routing, network-to-desktop Computer Telephony Integration (CTI), and multi-channel contact management to contact center agents over an IP network. It combines software IP automatic call distribution (ACD) functionality with Cisco Unified Communications in a unified solution that enables companies to rapidly deploy an advanced, distributed contact center infrastructure.

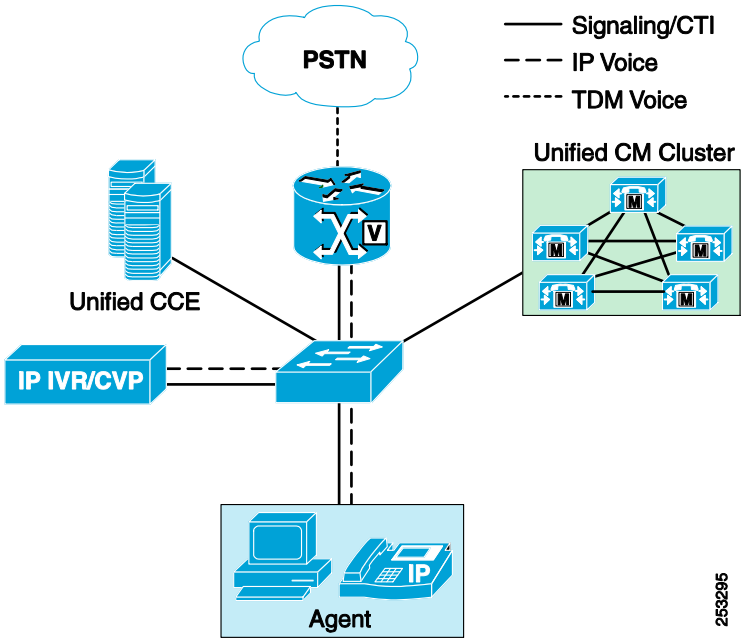
The reader of this document is expected to be familiar with the Unified CCE solution architecture and functionality as described in the *Installation and Configuration Guide for Cisco Unified Contact Center Enterprise & Hosted*, available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html. Make sure you become familiar with the concepts described in that manual for topics such as routes, labels, and dialed numbers.

This design guide describes the deployment models and their implications, including scalability, fault tolerance, and interaction between the solution components.

The Unified CCE product integrates with Cisco Unified Communications Manager (Unified CM), Cisco IP Interactive Voice Response (Unified IP IVR), Cisco Unified Customer Voice Portal (Unified CVP), Cisco Voice over IP (VoIP) Gateways and Cisco Unified IP Phones. Together these products provide Cisco Unified Communications and contact center solutions to achieve intelligent call routing, multi-channel automatic call distribution (ACD) functionality, interactive voice response (IVR), network call queuing, and consolidated enterprise-wide reporting. Unified CCE can optionally integrate with Cisco Unified ICM to support networking with legacy ACD systems while providing a smooth migration path to a converged communications platform.

The Unified CCE solution is designed for implementation in both single-site and multi-site contact centers. It utilizes your existing Cisco IP network to lower administrative expenses and extend the boundaries of the contact center enterprise to include branch offices, home agents, and knowledge workers. Figure 1-1 illustrates a typical Unified CCE setup.

Figure 1-1 Typical Unified CCE Solution Deployment



The Unified CCE solution consists of four primary Cisco software products:

- Unified Communications infrastructure: Cisco Unified Communications Manager (Unified CM)
- Queuing and self-service: Cisco Unified IP Interactive Voice Response (Unified IP IVR) or Cisco Unified Customer Voice Portal (Unified CVP)
- Contact center routing and agent management: Unified CCE. The major components are CallRouter, Logger, Peripheral Gateway, Administration & Data Server/Administration Client.
- Agent desktop software: Cisco Agent Desktop (CAD), Cisco Toolkit Agent Desktop (CTI OS), or integrations with third-party customer relationship management (CRM) software through Cisco Unified CRM Connector.

The solution is built on the Cisco IP Telephony infrastructure, which includes:

- Cisco Unified IP Phones
- Cisco voice gateways
- Cisco LAN/WAN infrastructure

The following sections discuss each of the software products in more detail and describe the data communications between each of those products. For more information on a particular product, refer to the specific product documentation available online at: http://www.cisco.com

Solution Components

Cisco Unified Communications Manager

Cisco Unified Communications Manager (Unified CM) is a software application that controls the voice gateways and IP phones, thereby providing the foundation for a Voice over IP (VoIP) solution. Unified CM runs on Cisco Media Convergence Servers (MCS). The software running on a server is referred to as a Unified CM server. Multiple Unified CM servers can be grouped into a cluster to provide for scalability and fault tolerance. Unified CM communicates with the gateways using standard protocols

such as H.323, Media Gateway Control Protocol (MGCP), and Session Initiation Protocol (SIP). Unified CM communicates with the IP phones using SIP or Skinny Call Control Protocol (SCCP). For details on Unified CM call processing capabilities and clustering options, refer to the latest version of the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at:

http://www.cisco.com/go/ucsrnd

Unified CM communicates with Unified CCE via the Java Telephony Application Programming Interface (JTAPI). A single Unified CM subscriber server is capable of supporting hundreds of agents. In a fault-tolerant design, a Unified CM cluster is capable of supporting thousands of agents. However, the number of agents and the number of busy hour call attempts (BHCA) supported within a cluster varies and must be sized according to guidelines defined in "Chapter 11, "Sizing Cisco Unified Communications Manager Servers".

Typically, when designing a Unified CCE solution, you first define the deployment scenario, including the arrival point(s) for voice traffic and the location(s) of the contact center agents. After defining the deployment scenario, you can determine the sizing of the individual components within the Unified CCE design, including how many Unified CM servers are needed within a Unified CM cluster, how many voice gateways are needed for each site and for the entire enterprise, how many servers and what types of servers are required for the Unified CCE software, and how many Unified IP IVR or Unified CVP servers are needed.

Cisco Voice Gateways

When you select voice gateways for a Unified CCE deployment, it is important to select voice gateways that satisfy not only the number of required PSTN trunks but also the busy hour call completion rate on those trunks. Busy hour call completion rates per PSTN trunk are typically higher in a contact center than in a normal office environment. For Cisco Catalyst Communications Media Module (CMM) voice gateways used in pure contact center deployments, provision a maximum of four T1/E1 interfaces to ensure that the call processing capacity of the voice gateway is satisfactory.

Agent Phones

Prior to release 8.0, Unified CCE supported monitoring of only a single line for all agent devices (Single Line Agent Mode).

Unified CCE Release 8.0 added support for monitoring multiple agent lines when Multi Line Agent Mode is enabled for the Peripheral. This feature provides the following capabilities:

- Monitoring and reporting of calls on all lines on the phone. For additional details on reporting in a multi-line environment, refer to the *Release Notes for Cisco Unified ICM/Contact Center Enterprise & Hosted, Release 8.0(1)* available at:

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html

- Other than placing a call, all other call control on the non-ACD extensions is supported from Multi Line capable desk tops, except for call initiation. Calls initiated from the hard phone can be controlled after initial call setup.
- Allows Unified CCE to support Join Across Line (JAL) and Direct Transfer Across Line (DTAL) features on the phone. These phone features are supported in all of our supported phone families.
- Requires a busy trigger of 1 (i.e. no call waiting), although calls can be forwarded to other extensions on the phone when busy.
- Requires a maximum of 2 call appearances.
- Supports a maximum of 4 lines per phone, one ACD line and up to three non-ACD lines.
- Shared lines are not supported on ACD and non-ACD lines.

- Call Park is not supported on ACD and non-ACD lines.
- Unified CCE may not be backward compatible with third-party CTI Applications when Multi Line Agent Mode is enabled. Multi-line support must be validated with the third-party vendor.

For a list of supported agent phones, refer to the *Cisco Unified Contact Center Enterprise (Unified CCE) Software Compatibility Guide*, available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_device_support_tables_list.html.

There are three families of phones in the Cisco portfolio:

Cisco Unified IP Phones 7900 Series

- The Unified CCE Agent Phone device supported prior to release 8.0 of Unified CCE.
- Join Across Line (JAL) and Direct Transfer Across Line (DTAL) are disabled by default.

Cisco Unified IP Phones 6900 Series

- Outbound campaign capability requires Cisco Unified Contact Center Enterprise 7.5(6) or later release.
- The 6900 Series phones do not support call waiting.
- Join Across Line (JAL) and Direct Transfer Across Line (DTAL) features are enabled by default.
- The 6900 Series phones are supported in Single Line mode only when JAL and DTAL are disabled.
- Unified CM silent monitoring and recording and Remote Silent Monitoring (RSM) is supported with the 8.5.(4) firmware load or higher.

Cisco Unified IP Phones 8900 Series and 9900 Series

- The 8900 Series and 9900 Series phones support cancel and swap features.
- Join Across Line (JAL) and Direct Transfer Across Line (DTAL) features are on all of the time, so contact center agents can merge or transfer calls on the ACD line with calls on other extensions on the phone. Because of this, 8900 Series and 9900 Series phones are supported only when the Multi Line Agent Mode feature of Unified CCE is enabled.

Cisco Unified Customer Voice Portal (Unified CVP)

Unified CVP is a software application running on industry standard servers such as Cisco Media Convergence Servers (MCS). It provides prompting, collecting, queuing, and call control services using standard web-based technologies. The Unified CVP architecture is distributed, fault tolerant, and highly scalable. With the Unified CVP system, voice is terminated on Cisco IOS gateways that interact with the Unified CVP application server using VoiceXML (speech) and H.323 or SIP (call control).

The Unified CVP software is tightly integrated with the Cisco Unified CCE software for application control. It interfaces with Unified CCE using the VRU Peripheral Gateway Interface. The Unified CCE scripting environment controls the execution of building-block functions such as play media, play data, menu, and collect information. The Unified CCE script can also invoke external VoiceXML applications to be executed by the Unified CVP VoiceXML Server, an Eclipse and J2EE- based scripting and web server environment. VoiceXML Server is well suited for sophisticated and high-volume IVR applications, and it can interact with custom or third-party J2EE-based services. These applications can return results and control to the Unified CCE script when complete. Advanced load balancing across all Unified CVP solution components can be achieved by Cisco Content Services Switch (CSS) and Cisco IOS Gatekeepers or Cisco Unified Presence SIP Proxy Servers.

Unified CVP can support multiple grammars for prerecorded announcements in several languages. Unified CVP can optionally provide automatic speech recognition and text-to-speech capability. Unified CVP can also access customer databases and applications via the Cisco Unified CCE software.

Unified CVP also provides a queuing platform for the Unified CCE solution. Voice and video calls can remain queued on Unified CVP until they are routed to a contact center agent (or external system). The system can play back music or videos while the caller is on hold; and when Unified CCE routes the call to an agent, he or she is able to push videos to a caller from the agent desktop application. For more information, refer to the latest version of the *Cisco Unified Customer Voice Portal (CVP) Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd.html

Cisco Unified IP Interactive Voice Response (Unified IP IVR)

The Unified IP IVR provides prompting, collecting, and queuing capability for the Unified CCE solution. Unified IP IVR does not provide call control as Unified CVP does because it is behind Unified CM and under the control of the Unified CCE software via the Service Control Interface (SCI). When an agent becomes available, the Unified CCE software instructs the Unified IP IVR to transfer the call to the selected agent phone. The Unified IP IVR then requests Unified CM to transfer the call to the selected agent phone.

Unified IP IVR is a software application that runs on Cisco MCS Servers. You can deploy multiple Unified IP IVR servers with a single Unified CM cluster under control of Unified CCE.

Unified IP IVR has no physical telephony trunks or interfaces like a traditional IVR. The telephony trunks are terminated at the voice gateway. Unified CM provides the call processing and switching to set up a G.711 or G.729 Real-Time Transport Protocol (RTP) stream from the voice gateway to the Unified IP IVR. The Unified IP IVR communicates with Unified CM via the Java Telephony Application Programming Interface (JTAPI), and the Unified IP IVR communicates with Unified CCE via the Service Control Interface (SCI) with a VRU Peripheral Gateway or System Peripheral Gateway.

Chapter 9, “Sizing Contact Center Resources ” discusses how to determine the number of IVR ports required. For deployments requiring complete fault tolerance, a minimum of two Unified IP IVRs is required. see Chapter 3, “Design Considerations for High Availability” which provides details on Unified CCE fault tolerance.

Unified Presence Server

Cisco Unified Presence Server is used by the Cisco Agent Desktop and Cisco Unified Expert Advisor products in the solution to locate appropriate resources (agents or other employees of the enterprise) for managing the call. See Chapter 4, “Unified Contact Center Enterprise Desktop”for details on Cisco Agent Desktop andChapter 7, “Cisco Unified Expert Advisor Option” for details on Unified Expert Advisor.

Unified ICM

Unified CCE may be deployed with Unified ICM to form a complete enterprise call management system. Unified ICM interfaces with ACDs from various vendors (including Cisco Unified CCE and Unified Expert Advisor), VRUs (including Cisco Unified IP-IVR and Unified CVP) and telephony network systems to efficiently and effectively treat customer contacts such as calls and emails, where the resources are located anywhere in the enterprise.

Unified CCE may be deployed in a "hybrid" model with Unified ICM, where the CallRouter, Logger, Administrative & Data Servers, and other components are shared between Unified ICM and Unified CCE. (See Unified CCE Software Components for a description of these components).

Alternatively, Unified CCE may be deployed in a "parent/child" model where Unified ICM is the parent, and Unified CCE is the child. This closely resembles the deployment model of Unified ICM with ACDs from other vendors. It is used for a highly scalable deployment because it provides CallRouters, data servers, and so forth for each product, although there are more components to manage and maintain. It is also used for a distributed model where isolation is needed between Unified ICM and Unified CCE, such as in an outsourced operation.

Unified Expert Advisor

Cisco Unified Expert Advisor (Unified Expert Advisor) is an optional component in a Unified ICM and Unified CCE deployments. When deployed with Unified ICM, Unified Expert Advisor is very much like other Unified ICM ACD integrations. It has its own type of PG and agents that are known as *experts* or *expert advisors,* who are part of the enterprise but usually not a part of the call center. The expert advisor has an expertise that may be tapped by traditional agents or tapped directly by callers into the contact center. Unified ICM skill groups can be created in order to route calls to those agents. Agent availability and status changes are also tracked and reported on, just as they would be for a typical ACD peripheral.

When deployed with Unified CCE, Unified Expert Advisor is treated as a hybrid integration, where the Unified CCE Routing engine routes to Agents associated with Unified CCE and Skill Group targets associated with *expert advisors*.

For more information on Unified Expert Advisor, see Chapter 7, "Cisco Unified Expert Advisor Option".

Cisco Unified Contact Center Enterprise (Unified CCE)

Cisco Unified CCE is the software application that provides the contact center features, including agent state management, agent selection, call routing and queue control, IVR control, CTI Desktop screen pops, and contact center reporting. Unified Contact Center Enterprise (Unified CCE) runs on Cisco MCS servers or exact equivalents, unless otherwise specified in Chapter 10, "Sizing Unified CCE Components and Servers" and the *Hardware & System Software Specification (Bill of Materials) for Cisco Unified ICM/Contact Center Enterprise & Hosted, Release 8.0(1)* available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_technical_reference_list.html. It relies on the Microsoft Windows 2003 operating system software and Microsoft SQL Server 2005 database management system. The supported servers can be single, dual, or quad Pentium CPU servers in single or multi-core variations with varying amounts of RAM. This variety of supported servers allows the Unified CCE software to scale and to be sized to meet the needs of the deployment requirements. (Chapter 10, "Sizing Unified CCE Components and Servers" provides details on server sizing.)

Unified CCE Software Components

This section describes the main components of the Unified CCE Product. Following sections describe some key concepts and terminology, and go into more detail on some of the components.

The Cisco Unified CCE software is a collection of components that can run on multiple servers. The number and type of components that can run on one server is primarily based upon busy hour call attempts (BHCA) and the size of the server being used (single, dual, or quad CPU). Other factors that

impact the hardware sizing are the number of agents, the number of skill groups per agent, the number of Unified IP IVR ports, the number of VRU Script nodes in the routing script, Extended Call Context (ECC) usage, and which statistics the agents need at their desktops.

The core Unified CCE software components are listed here and described in greater detail later in this chapter:

Unified CCE Software Components	Description
CallRouter	Makes all routing decisions on how to route a call or customer contact. Often just referred to as “Router” in the context of Unified CCE components.
Logger	The database server that stores contact center configuration and temporarily stores historical reporting data for distribution to the data servers
CTI Object Server (CTI OS)	CTI interface for Agent Desktops.
Peripheral Gateway (PG)	Interfaces to various ‘peripheral’ devices, specifically to Unified CM, VRU (Unified IP IVR and/or Unified CVP), or Multichannel products (EIM/WIM for email and chat). The PG includes one or more Peripheral Interface Managers (PIMs) for the specific device interfaces.
Agent PG	PG that has a Unified CM PIM.
Unified CM Peripheral Interface Manager (PIM)	Part of a PG that interfaces to a Unified CM cluster via the JTAPI protocol.
VRU PIM	Part of a PG that interfaces to the Unified IP IVR or Unified CVP via the Service Control Interface (SCI) protocol.
MR PIM	Part of a PG that interfaces to call center Multimedia products, specifically EIM and WIM for email and chat.
CTI Server	Part of the PG that interfaces to CTI OS and provides an open CTI interface, which is useful for some server-to-server communications.
Network Interface Controller (NIC)	Interfaces to the public switched telephone network (PSTN), which enables Unified CCE to control how the PSTN routes a call.
Administration & Data Server	Configuration interface and real-time and historical data storage (for example, for reporting). There are several different deployment models described later in this chapter.
Administration Client	Configuration interface. This component has a subset of the functionality of the Administration & Data Server. It is a “lighter weight” deployment because it does not require a local database, and it is deployed to allow more places from which to configure the solution.
Cisco Unified Intelligence Center (Unified IC)	Provides Web browser-based real-time and historical reporting. Unified IC also works with other Cisco Unified Communications products.
WebView Reporting Server	Legacy component that provides Web browser-based real-time and historical reporting.

The combination of CallRouter and Logger is called the Central Controller. When the CallRouter and Logger modules run on the same server, the server is referred to as a *Rogger*. When the CallRouter, Logger, and Peripheral Gateway modules run on the same server, the server is referred to as a *Progger*. In lab environments, the system Administration & Data Server can also be loaded onto the Progger to create a server known as a *Sprawler* configuration; however, this configuration is approved only for lab use and is not supported in customer production environments.

Redundancy and Fault Tolerance

The CallRouter and Logger are deployed in a paired redundant fashion. This redundant configuration is also referred to as *duplex mode*, whereas a non-redundant configuration is said to be running in *simplex mode*. (**Note**: Simplex mode is *not* supported for production environments). The two sides of the redundant deployment are referred to as side A and side B. For example, CallRouter A and CallRouter B are redundant instances of the CallRouter running on two different servers. When both sides are running (normal operation), it is referred to as *duplex* mode. When one side is down, it is said to be running in *simplex* mode. The two sides are for redundancy, not load-balancing. Either side is capable of running the full load of the solution. The A and B sides are both executing the same set of messages and, therefore, producing the same result. In this configuration, logically, there appears to be only one CallRouter. The CallRouters run in synchronized execution across the two servers, which means both sides of the duplex servers process every call. In the event of a failure, the surviving CallRouter will pick up the call mid-stream and continue processing in real-time and without user intervention.

The Peripheral Gateway components run in hot-standby mode, meaning that only one of the Peripheral Gateways is actually active and controlling Unified CM or the IVR. When the active side fails, the surviving side automatically takes over processing of the application. During a failure, the surviving side is running in simplex mode and will continue to function this way until the redundant side is restored to service, when it will automatically return to duplex operation.

The CTI OS component provides fault tolerance through a pair of servers that operate together and back up each other. There is no notion of an active and passive server, or of a primary and secondary server. Both servers are always active. Clients may connect to either server. In the event of the failure of any one server, clients can automatically reconnect to the alternate server.

The Administration & Data Servers for configuration and real-time data are deployed in pairs for fault tolerance, with multiple pairs deployed for scalability. The data flows are described in the detailed section on Administration & Data Server/Administration Client.

The Administration & Data Servers for historical data follow an n+1 architecture for redundancy and scalability, with each choosing a Logger side (A or B) as its preferred and primary data source.

WebView servers are deployed in an n+1 architecture for redundancy and scalability. They can be co-resident with the Administration & Data Server for historical data, or deployed in standalone web-server mode to achieve higher scalability in terms of reporting users that need access to the application for real-time and historical reporting. (See Chapter 10, “Sizing Unified CCE Components and Servers” for more details.)

Customer Instance and Unified CCH

The Cisco Unified Contact Center Hosted (Unified CCH) solution is largely the same as Unified CCE, but it supports multi-tenant or shared servers to manage multiple *customer instances*.

All Unified CCE systems are deployed as a single instance (using the same instance name and number in setup) across all the Unified CCE components.

Peripheral Gateway (PG) and PIMs

For each Unified CM cluster in the Unified CCE environment, there is a Unified CM PIM on an Agent Peripheral Gateway. For scalability requirements, some deployments may require multiple PIMs for the same Unified CM cluster; they may be on the same PG and physical server or they may be separate.

For each Agent PG, there is one CTI Server component and one or more CTI OS components to communicate with the desktops associated with the phones for that Unified CM cluster.

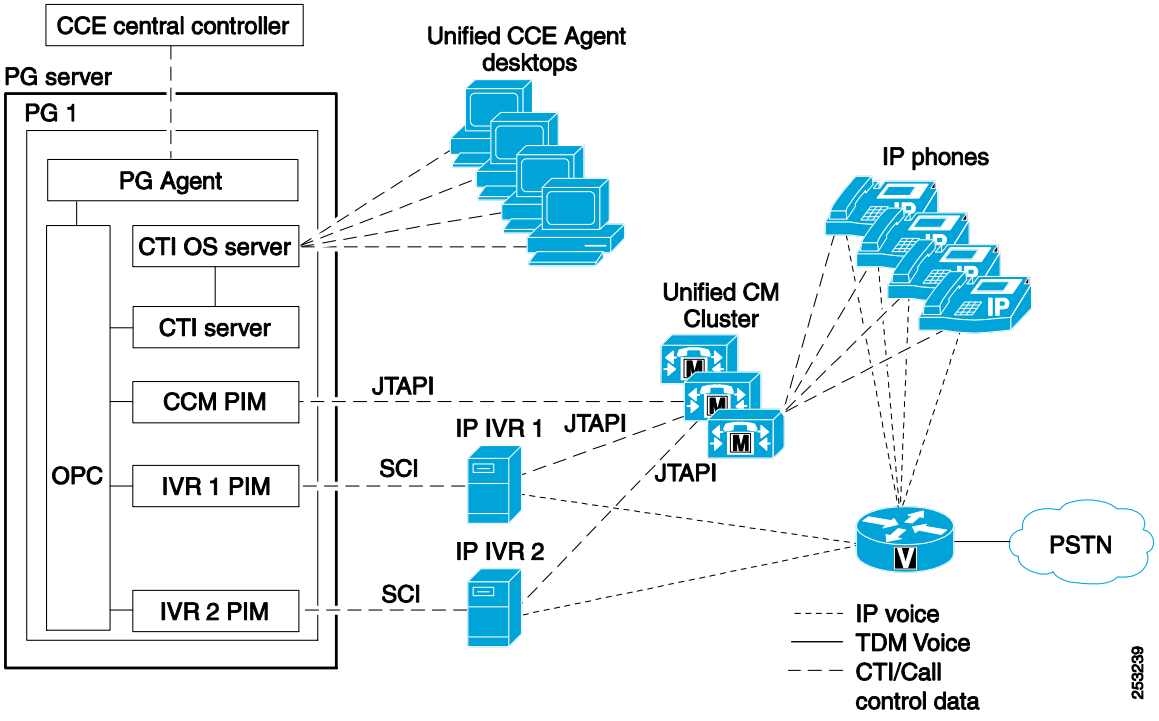
Note The CTI OS components on Side A and Side B are simultaneously active to load-balance desktop communication.

For each Unified IP IVR or CVP Call Server, there is one VRU PIM. VRU PIMs may be part of the Agent PG.

Often, the Unified CM PIM, the CTI Server, the CTI OS, and multiple VRU PIMs may run on the same server.

Internal to the PG is a process called the PG Agent, which communicates to the Central Controller. Another internal PG process is the Open Peripheral Controller (OPC), which enables the other processes to communicate with each other and is also involved in synchronizing PGs in redundant PG deployments. Figure 1-2 shows the communications among the various PG software processes.

Figure 1-2 Communications Among Peripheral Gateway Software Processes



In larger, multi-site (multi-cluster) environments, multiple Agent PGs are usually deployed. When multiple Unified CM clusters are deployed, Unified CCE tracks all the agents and calls centrally, and is able to route the calls to the most appropriate agent independent of the site or cluster that they are using, thus making them all appear to be part of one logical enterprise-wide contact center with one enterprise-wide queue.

Network Interface Controller (NIC)

The Network Interface Controller (NIC) is an optional component that interfaces to the public switched telephone network (PSTN). Intelligently routing a call before the call is delivered to any customer premise equipment is referred to as pre-routing. Only certain PSTNs have NICs supported by Unified CCE. For a detailed list of PSTN NICs and details on Unified CCE pre-routing, refer to the *Pre-installation Planning Guide for Cisco Unified ICM Enterprise & Hosted*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1001/prod_installation_guides_list.html

Unified CCE Agent Desktop Options

See Chapter 4, "Unified Contact Center Enterprise Desktop" for detailed information on the options for Agent Desktops, including CAD and CTI OS interfaces

Cisco offers the following interfaces for Unified CCE agents (see Figure 1-3):

- Cisco Agent Desktop

  Cisco Agent Desktop provides an out-of-the-box, feature-rich desktop solution for Unified CCE. The desktop application can be deployed in various ways:

  - Windows application
  - Browser-based application

    Cisco Unified IP Phone Agent, where there is no desktop application at all but just an XML application on the IP phone
  - Cisco CTI Desktop Toolkit

    The CTI Desktop Toolkit provides a software toolkit for building custom desktops, desktop integrations into third-party applications, or server-to-server integrations to third-party applications.

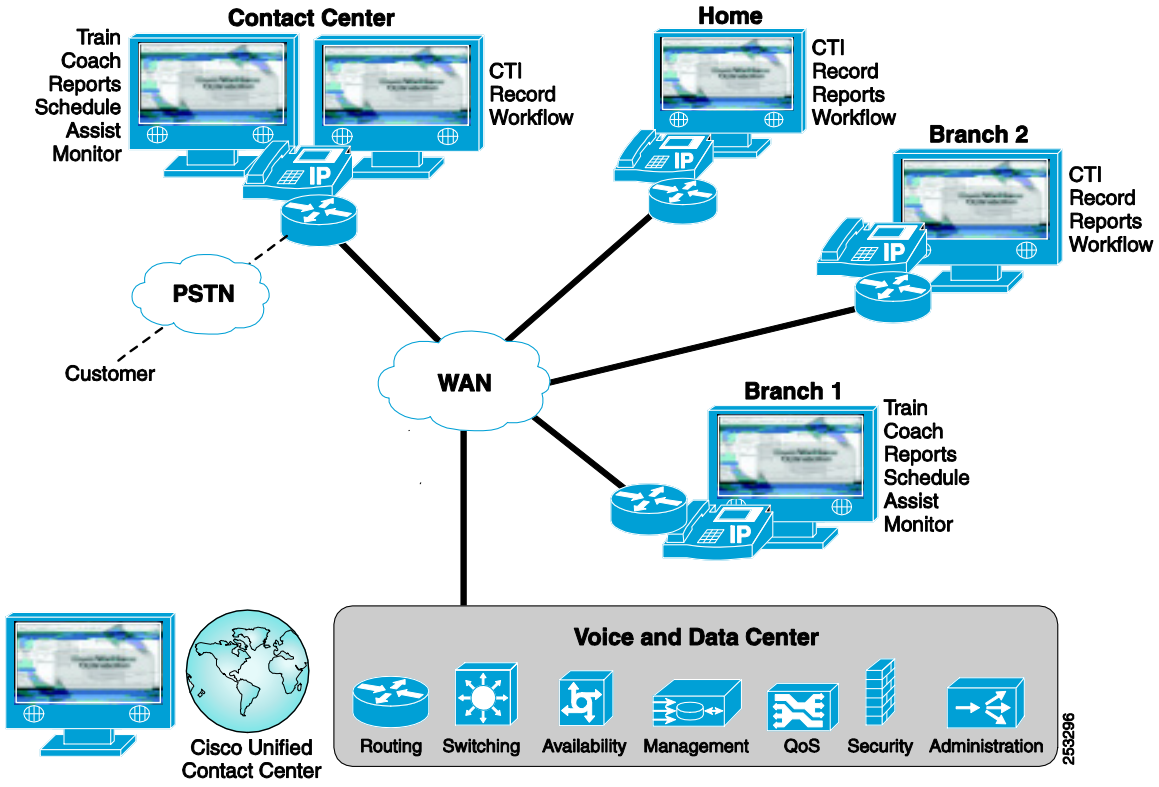
- CRM Connectors

  Cisco offers pre-built, certified CRM Connectors for CRM packages, including SAP, Siebel (using CTI OS driver), Salesforce.com, Microsoft Dynamics CRM, and Peoplesoft. These integrated solutions enable call control from the CRM user interface (Answer, Drop, Hold, Un-Hold, Blind or Warm Transfers, and Conferences), outbound and consultative calls from the CRM desktop, and delivery and manipulation of Call Context Data (CTI screen pop).

  Agents using a third-party CRM user interface connected through a CRM Connector can be supervised using a CTI Desktop Toolkit-based supervisor desktop.

For more information about desktop selection and design considerations, see Chapter 4, "Unified Contact Center Enterprise Desktop".

Figure 1-3 Variety of Agent Interfaces for Unified CCE



Administration & Data Server/Administration Client

Administration & Data Servers have several roles: Administration, Real-time data server, Historical Data Server, and Detail Data Server. A Unified CCE deployment must have Administration & Data Servers to fill these roles. The servers may be deployed in the following combinations to achieve the needed scalability with the minimum number of servers:

- Administration Server and Real-time Data Server (AW)
- Configuration-Only Administration Server
- Administration Server, Real-time and Historical Data Server, and Detail Data Server (AW-HDS-DDS)
- Administration Server and Real-time and Historical Data Server (AW-HDS)
- Historical Data Server and Detail Data Server (HDS-DDS)

Note See Chapter 2, "Deployment Models" for more details on deployment options and requirements.

- An Administration Client (formerly known as a "client AW") serves the administration role but is deployed as a client to an Administration Server for scalability. The Administration Client may view and modify the configuration, and receive real-time reporting data from the AW, but it does not store the data itself and does not have a database.

Each Administration & Data Server must be installed on a separate server for production systems to ensure no interruptions to the real-time call processing of the CallRouter and Logger processes. For lab or prototype systems, the Administration & Data Server (with the WebView Server option) can be installed on the same server as the CallRouter and Logger.

Administration Server and Administration Client

The Administration Server, Configuration-Only Administration Server, and Administration Client provide a Configuration Manager tool used to configure Unified CCE. This includes, for example, the ability to add agents, add skill groups, assign agents to skill groups, add dialed numbers, add call types, assign dialed numbers to call types, or assign call types to routing scripts.

The Administration Server and Administration Client also have a tool "Script Editor" which is used to build routing scripts. Routing scripts specify how to route and queue a contact (that is, the script identifies which skill group or agent will handle a particular contact).

- The Administration Server and Configuration-Only Administration Server also support the following configuration tools:
- Agent Re-skilling Web Tool (Unified CCE only)
- Configuration Management service (CMS) Node.
- Internet Script Editor Server - HTTPS (default protocol) connection for Script Editor clients

For details on the use of these and other configuration tools, refer to the *Administration Guide for Cisco Unified ICM/Contact Center Enterprise & Hosted*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_maintenance_guides_list.html

The Administration Server is deployed as part of the "Administration and Real Time Data Server," known as AW. AWs are deployed in pairs for fault tolerance. During normal operation, the "primary AW" communicates directly with the Central Controller for configuration data (see Figure 1-4) and the "secondary AW" connects to the primary AW for the data. If the primary AW fails, the secondary AW connects to the central controller. Both types of AW store the configuration and real time data in the AW Database, or AWDB. Each AW can be deployed in the same location as, or remote from, the Central Controller. A secondary AW need not be co-located with the primary AW.

- Multiple Administration Clients can be deployed and connected to either primary or secondary AWs. An Administration Client must be geographically local to its AW.

ConfigurationOnly Administration Servers are the same as AWs, but without the real-time data. As such, Administration Clients cannot connect to them and they cannot display real-time data in Script Editor. They can be deployed in a multi-instance configuration for Hosted CCE (seeFigure 1-5).

Figure 1-4 Communication Between Unified CCE Central Controller and Administration & Data Server

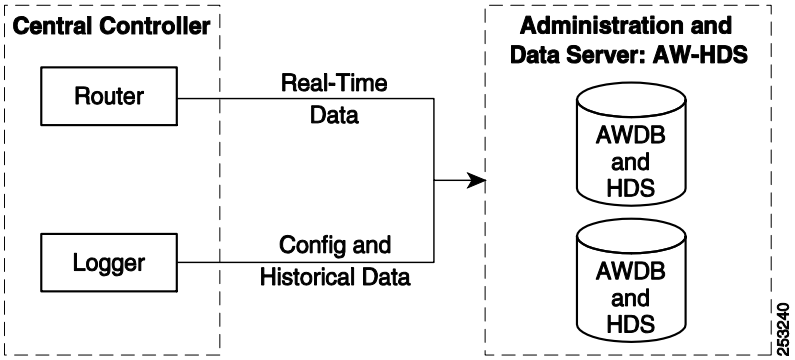
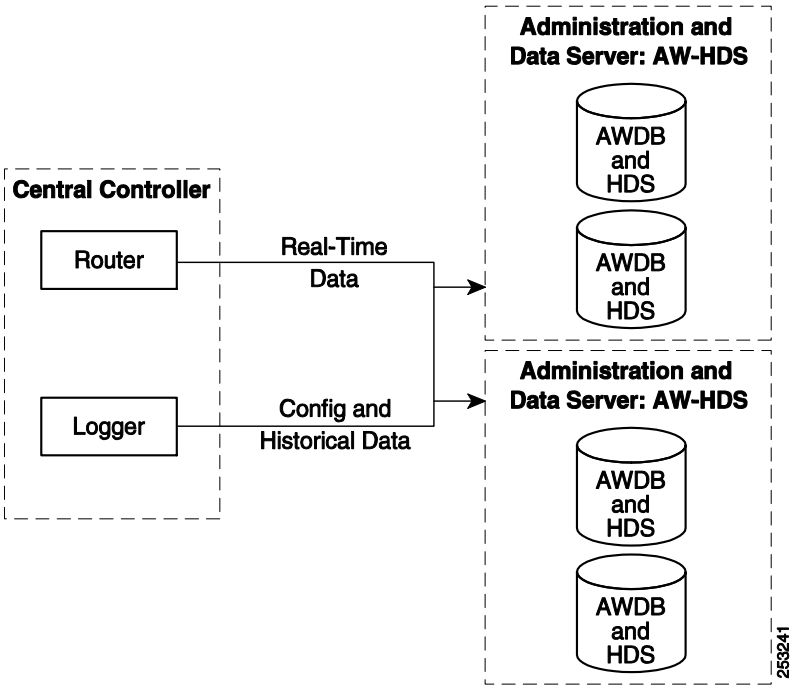


Figure 1-5 Communication Between Unified CCE Central Controller and Multiple Administration & Data Servers



AWs, Configuration-Only Administration Servers, and Administration Clients may operate only as a single instance on a given server. In a hosted environment, multiple instances may be installed and configured, and the "Select Administration Instance" tool may be used to switch between the instances.

Real Time Data Server

The Real-Time Data Server portion of the AW uses the AW database to store real-time data along with the configuration data. Real-time reports combine these two types of data to present a near-current transient snapshot of the system.

Historical Data Server (HDS) and Detail Data Server (DDS)

The Historical Data Server (HDS) and Detail Data Server (DDS) are used for longer-term historical data storage. The HDS stores historical data summarized in 15 or 30 minute intervals, and is used for reporting. DDS stores detailed information about each call or call segment, and is used for call tracing. Data may be extracted from either of these sources for warehousing and custom reporting.

Typically these Data Servers are deployed with a primary AW as a single server serving all three roles (AW-HDS-DDS). In very large deployments it might be desirable to separate them for scalability.

Unified CCE Reporting

The Unified CCE Reporting solution provides an interface to access data describing the historical and real-time states of the system.

The reporting solution consists of the following components:

- Unified IC or WebView — reporting user interfaces
- Configuration and Reporting Data — contained on Administration & Data Server(s)

Note Reporting concepts and data descriptions are described in the *Reporting Guide for Cisco Unified ICM/Contact Center Enterprise & Hosted* available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_configuration_examples_list.html. (This description is independent of the reporting user interface being used.)

Cisco Unified Intelligence Center

Cisco Unified Intelligence Center (Unified IC) is an advanced reporting product used for Unified CCE and other products. This platform is a web-based application offering many Web 2.0 features, high scalability, performance, and advanced features such as the ability to integrate data from other Cisco Unified Communications products or third-party data sources. Unified IC incorporates a security model that defines different access and capabilities for specific users.

Unified IC Standard is included with Unified CCE. Unified IC Premium is an optional product with additional features. Refer to the *Cisco Unified Intelligence Suite Intelligence Center User Guide*, available at http://www.cisco.com/en/US/products/ps9755/products_user_guide_list.html. Unified IC must be installed on a separate server; it cannot be co-resident with other Unified CCE components.

WebView

WebView is a web-based application for Unified CCE reporting. WebView is being replaced by Unified IC. In this release, both products are available to allow users to smoothly transition to Unified IC. The two products may be used concurrently, subject to the scalability of that deployment as described in Chapter 10, "Sizing Unified CCE Components and Servers".

WebView performs the basic operations of gathering user input, querying the databases, and presenting the requested data for both real-time and historical data.

Note WebView does not support historical data that is collected in 15 minute intervals (a feature new in Unified CCE 8.0). WebView supports historical reporting only for data collected in half-hour intervals. WebView also does not supply reports for the call type or skill group data (also new in Unified CCE 8.0). Unified IC must be used for 15-minute data intervals or call type and skill group reporting.

WebView has the ability to export data, launch scheduled reports, save and share report settings, and mark favorite reports. It also has features to display service-affecting events reported by the system.

WebView can be installed on an Administration & Data Server or, to increase scalability, it can be installed on a standalone server. For information on the reporting deployment options, see Chapter 10, "Sizing Unified CCE Components and Servers" and Chapter 8, "Securing Cisco Unified Contact Center Enterprise".

The WebView architecture is described in the *WebView Installation and Administration Guide*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html

For a description of all of the reports provided with WebView, refer to the *WebView Template Reference Guide for Cisco Unified Contact Center Enterprise & Hosted*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html

Unified Contact Center Management Portal (CCMP)

The Unified Contact Center Management Portal provides a simple to use web-based user interface to streamline the day-to-day provisioning and configuration operations performed by a contact center manager, team lead, or administrator. The Management Portal provides the following key benefits:

- Simple to use web user interface for performing basic tasks such as move/add/modify phones, agents, skill groups, teams, and other common contact center administrative functions for an IP contact center
- Unified Configuration; that is, tenant provisioning of both the applicable IP contact center elements and the Cisco Unified Communications Manager components through a single task-based web interface
- Partitioned System supporting multiple business units with complete autonomy
- Hierarchical Administration supporting multiple business-level users, where each user is defined with specific roles and responsibilities
- Audit Trail Reports that detail configuration changes and usage by all users of the management portal

See Chapter 13, "Cisco Unified Contact Center Management Portal" for more information on Cisco Unified Contact Center Management Portal.

JTAPI Communications

In order for JTAPI communications to occur between Unified CM and external applications such as Unified CCE and Unified IP IVR, a JTAPI user ID and password must be configured within Unified CM. Upon startup of the Unified CM PIM or upon startup of the Unified IP IVR, the JTAPI user ID and password are used to log in to Unified CM. This login process by the application (Unified CM PIM or Unified IP IVR) establishes the JTAPI communications between the Unified CM cluster and the application. A single JTAPI user ID is used for all communications between the entire Unified CM cluster and Unified CCE. A separate JTAPI user ID is also required for each Unified IP IVR server. In a Unified CCE deployment with one Unified CM cluster and two Unified IP IVRs, three JTAPI user IDs are required: one JTAPI user ID for Unified CCE and two JTAPI user IDs for the two Unified IP IVRs.

The Unified CM software includes a module called the CTI Manager, which is the layer of software that communicates via JTAPI to applications such as Unified CCE and Unified IP IVR. Every node within a cluster can execute an instance of the CTI Manager process, but the Unified CM PIM on the PG communicates with only one CTI Manager (and thus one node) in the Unified CM cluster. The CTI

Manager process communicates CTI messages to/from other nodes within the cluster. For example, suppose a deployment has a voice gateway homed to node 1 in a cluster, and node 2 executes the CTI Manager process that communicates to Unified CCE. When a new call arrives at this voice gateway and needs to be routed by Unified CCE, node 1 sends an intra-cluster message to node 2, which will send a route request to Unified CCE to determine how the call will be routed.

Each Unified IP IVR also communicates with only one CTI Manager (or node) within the cluster. The Unified CM PIM and the two Unified IP IVRs from the previous example could each communicate with different CTI Managers (nodes) or they could all communicate with the same CTI Manager (node). However, each communication uses a different user ID. The user ID is how the CTI Manager keeps track of the different applications.

When the Unified CM PIM is redundant, only one side is active and in communication with the Unified CM cluster. Side A of the Unified CM PIM communicates with the CTI Manager on one Unified CM node, and side B of the Unified CM PIM communicates with the CTI Manager on another Unified CM node. The Unified IP IVR does not have a redundant side, but the Unified IP IVR does have the ability to fail over to another CTI Manager (node) within the cluster if its primary CTI Manager is out of service. (For more information about failover, see Chapter 3, "Design Considerations for High Availability".

The JTAPI communications between the Unified CM and Unified CCE include three distinct types of messaging:

- Routing control

Routing control messages provide a way for Unified CM to request routing instructions from Unified CCE.

- Device and call monitoring

Device monitoring messages provide a way for Unified CM to notify Unified CCE about state changes of a device (phone) or a call.

- Device and call control

Device control messages provide a way for Unified CM to receive instructions from Unified CCE on how to control a device (phone) or a call.

A typical Unified CCE call includes all three types of JTAPI communications within a few seconds. When a new call arrives, Unified CM requests routing instructions from Unified CCE. For example, when Unified CM receives the routing response from Unified CCE, Unified CM attempts delivery of the call to the agent phone by instructing the phone to begin ringing. At that point, Unified CM notifies Unified CCE that the device (phone) has started ringing, and that notification enables the agent's answer button on the desktop application. When the agent clicks the answer button, Unified CCE instructs Unified CM to make the device (phone) go off-hook and answer the call.

In order for the routing control communication to occur, Unified CM requires the configuration of a CTI Route Point. A CTI Route Point is associated with a specific JTAPI user ID, and this association enables Unified CM to know which application provides routing control for that CTI Route Point. Directory (Dialed) Numbers (DNs) are then associated with the CTI Route Point. A DN is associated to a CTI Route Point that is associated with Unified CCE JTAPI user ID, and this enables Unified CM to generate a route request to Unified CCE when a new call to that DN arrives.

In order for the phones to be monitored and controlled, they also must be associated in Unified CM with a JTAPI user ID. In a Unified CCE environment, the IP phones are associated with Unified CCE JTAPI user IDs. When an agent logs in from the desktop, the Unified CM PIM requests Unified CM to allow the PIM to begin monitoring and controlling that phone. Until the login has occurred, Unified CM does not allow Unified CCE to monitor or control that phone. If the device has not been associated with a Unified CCE JTAPI user ID, then the agent login request will fail.

Support for Extension Mobility Cross Cluster (EMCC) is provided in Release 8.0. The Unified CCE PIM phone registers to the local Unified CM after Extension Mobility login, and it looks like an agent situated across a WAN. The Unified CCE peripheral is managing the agent devices based on Extension Mobility profile rather than on a phone device in the Application User on Unified CM. For more details, refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)*, available at:

http://www.cisco.com/go/ucsrnd.html

Because the Unified IP IVR also communicates with Unified CM using the same JTAPI protocol, these same three types of communications also occur with the Unified IP IVR. Unlike Unified CCE, the Unified IP IVR provides both the application itself and the devices to be monitored and controlled.

The devices that Unified CCE monitors and controls are the physical phones. The Unified IP IVR does not have real physical ports like a traditional IVR. Its ports are logical ports (independent software tasks or threads running on the Unified IP IVR application server) called CTI Ports. For each CTI Port on the Unified IP IVR, there needs to be a CTI Port device defined in Unified CM.

Unlike a traditional PBX or telephony switch, Unified CM does not select the Unified IP IVR port to which it will send the call. Instead, when a call needs to be made to a DN that is associated with a CTI Route Point that is associated with a Unified IP IVR JTAPI user, Unified CM asks the Unified IP IVR (via JTAPI routing control) which CTI Port (device) will handle the call. Assuming the Unified IP IVR has an available CTI Port, the Unified IP IVR will respond to the Unified CM routing control request with the Unified CM device identifier of the CTI Port that is going to handle that call.

When an available CTI Port is allocated to the call, a Unified IP IVR workflow is started within the Unified IP IVR. When the Unified IP IVR workflow executes the accept step, a JTAPI message is sent to Unified CM to answer the call on behalf of that CTI Port (device). When the Unified IP IVR workflow wants the call transferred or released, it again instructs Unified CM on what to do with that call. These scenarios are examples of device and call control performed by the Unified IP IVR.

When a caller releases the call while interacting with the Unified IP IVR, the voice gateway detects the caller release and notifies Unified CM via H.323 or Media Gateway Control Protocol (MGCP), which then notifies the Unified IP IVR via JTAPI. When DTMF tones are detected by the voice gateway, it notifies Unified CM via H.245 or MGCP, which then notifies the Unified IP IVR via JTAPI. These scenarios are examples of device and call monitoring performed by the Unified IP IVR.

In order for the CTI Port device control and monitoring to occur, the CTI Port devices on Unified CM must be associated with the appropriate Unified IP IVR JTAPI user ID. If you have two 150-port Unified IP IVRs, you would have 300 CTI ports. Half of the CTI ports (150) would be associated with JTAPI user Unified IP IVR #1, and the other 150 CTI ports would be associated with JTAPI user Unified IP IVR #2.

While Unified CM can be configured to route calls to Unified IP IVRs on its own, routing of calls to the Unified IP IVRs in a Unified CCE environment will be done by Unified CCE (even if you have only one Unified IP IVR and all calls require an initial IVR treatment). Doing so will ensure proper Unified CCE reporting. For deployments with multiple Unified IP IVRs, this routing practice also allows Unified CCE to load-balance calls across the multiple Unified IP IVRs.

Multichannel Subsystems: EIM/WIM

Unified CCE has the capability to provide a multichannel contact center with E-mail Interaction Manager (EIM) and Web Interaction Manager (WIM).

For design information on these products, refer to the *Cisco Unified Web and E-Mail Interaction Manager Solution Reference Network Design (SRND) Guide for Unified Contact Center Enterprise, Hosted, and ICM*, available at

http://www.cisco.com/en/US/products/ps7236/products_implementation_design_guides_list.html

Cisco Unified Outbound Option

Agents can handle both inbound and outbound contacts, which helps in optimizing contact center resources. The Cisco Unified Outbound Option enables the multi-functional contact center to take advantage of Cisco Unified CCE enterprise management. Contact center managers in need of outbound campaign solutions can take advantage of the enterprise view that Cisco Unified CCE maintains over agent resources. (See Chapter 5, "Oubound Option for Cisco Unified Contact Center Enterprise and Hosted" for details.)

Cisco Unified Mobile Agent

Cisco Unified CCE provides the capability for an agent to use any PSTN phone and a quality high-speed data connection between the agent desktop and the CTI OS server. (For design guidance and considerations for implementing Cisco Unified Mobile Agent, see Chapter 6, "Cisco Unified Mobile Agent ".)

Unified System CCE 7.x

Note There is *no* 8.0 release for Unified System CCE; it will remain at Release 7.5. Users will be able to migrate their Unified System CCE 7.1, 7.2 or 7.5 system directly to Unified CCE 8.0 using the Unified CCE installer.

Cisco Unified System Contact Center Enterprise 7.x (Unified System CCE 7.x) is a deployment model that simplifies installation and configuration by using three predefined configurations for Unified CCE. Unified System CCE 7.x uses a single installer to simplify installation and configuration, and it provides web-based administration. Configuration of Unified System CCE 7.x is further simplified by removing Services, Translation Routes, Device Targets, Labels, and Sub Skill Groups

Unified System CCE 7.x provides fault tolerance through the duplex operation on the Central Controller and Agent/IVR Controller. Unified System CCE 7.x can connect to a parent Unified ICM, and the connection is made between the child Unified CCE System PG and the parent Gateway PG.

For further information about Unified System CCE 7.x, refer to the *Cisco Unified Contact Center Enterprise 7.0, 7.1, and 7.2 Solution Reference Network Design (SRND).*

Serviceability

Diagnostic Tools

Unified CCE has a built-in web-based (REST-like) interface for diagnostics called the Diagnostic Framework, which is resident on every Unified CCE server. The Anaysis Manager functionality integrated with the Unified Communications Manager Real-Time Monitoring Tool (RTMT) is provided as the client-side tool to collect diagnostic information from this diagnostic framework. In addition to the Analysis Manager, a command line interface – Unified System CLI tool - is available as well that allows a client to access the diagnostic framework on any Unified Communications server. (A user need not use Remote Desktop to gain access first to use the CLI.) The Analysis Manager is planned to replace Support Tools going forward as its interface is consistent across the Cisco Unified Communications solution.

Using the Analysis Manager, the administrator can connect to one or more Unified Commincations devices to set trace levels, collect trace and log files, and gather platform and application configuration data as well as version and license information. The Analysis Manager is the one tool that allows administrators to collect diagnostic information from all Cisco Unified Communications applications and devices.

The Analysis Manager offers local user and/or domain security for authentication and secure HTTP to protect data exchanged by it and the diagnostic framework.

For more information about the Unified CCE Diagnostic Framework (that runs on every Unified CCE server), refer to the *Serviceability Best Practices Guide for Cisco Unified ICM/Contact Center Enterprise & Hosted* at:
http://www.cisco.com/en/US/products/sw/custcosw/ps1001/products_installation_and_configuration_guides_list.html

Solution Call Trace

Also integrated with the Unified CM Real-Time Monitoring Tool is the Analyze Call Path feature. This tool enables the administrator to search for failed calls (based on simple filter parameters) and then trace those calls through each component of a Unified Communications solution. The Analyze Call Path tool will first identify call records for failed calls (for example, dropped or abandoned calls) in a particular application. The administrator may then select a particular call and have the tool collect all related call records (and traces) from all other Unified Communications components such as Unified CVP or Unified CM. Analysis of the call record collection can then reveal the exact location and nature of the call failure.

Additionally, the Analyze Call Path tool allows the administrator to search for call records based on other criteria such as Originating Date/Time, Calling Number, or Called Number, and to collect all associated call records for display and analysis.

For more information on Unified Communications Analyze Call Path capabilities in RTMT, refer to the *Analysis Manager User Guide* at:
http://www.cisco.com/en/US/docs/voice_ip_comm/uc_system/UC_Analysis_Manager801/Analysis_Manager_User_Guide801.pdf.

Support Tools

Cisco Support Tools is an application that contains a suite of utilities that allow you to manage and troubleshoot servers that run a broad range of Cisco Unified Communications software components. Through Support Tools, you can troubleshoot configuration and performance problems on these systems from any machine running a supported version of Windows and Internet Explorer on your network that can access the Support Tools Server. (Support Tools requires a separate server to act as both the web server and a repository for collected diagnostic data.)

Access to utilities in the Support Tools suite is through a browser-based interface - the Support Tools Dashboard - installed on the Support Tools Server. Levels of security control both access to the Dashboard and the ability to use specific tools once logged in. In low bandwidth conditions (for example, via dial-up access) or when Web browsing is otherwise impractical, many Support Tools utilities can also be accessed and run via the command line interface.

Network Management Tools

Unified CCE is manageable via the Simple Network Management Protocol (SNMP). Unified CCE devices have a built-in SNMP agent infrastructure that supports SNMP v1, v2c, and v3, and it exposes instrumentation defined by the CISCO-CONTACT-CENTER-APPS-MIB. This MIB provides configuration, discovery, and health instrumentation that can be monitored by standard SNMP management stations. Moreover, Unified CCE provides a rich set of SNMP notifications that will alert

administrators of any faults in the system. Unified CCE also provides a standard syslog event feed (conforming to RFC-3164) for those administrators who want to take advantage of a more verbose set of events.

For managing a Unified Communications deployment, customers are encouraged to use the Cisco Unified Operations Manager (Unified Operations Manager) product. Unified Operations Manager is a member of the Cisco Unified Communications family of products and provides a comprehensive and efficient solution for network management, provisioning, and monitoring of Cisco Unified Communications deployments.

Unified Operations Manager monitors and evaluates the current status of both the IP communications infrastructure and the underlying transport infrastructure in your network. Unified Operations Manager uses open interfaces such as Simple Network Management Protocol (SNMP) and Hypertext Transfer Protocol (HTTP) to remotely poll data from different devices in the IP communications deployment. In addition to the infrastructure, Unified Operations Manager offers capabilities specific to Unified Communications applications as well, including Unified CCE. For more information about Unified Operations Manager, refer to:
http://www.cisco.com/en/US/products/ps6535/tsd_products_support_series_home.html

For more information about configuring the Unified CCE SNMP agent infrastructure and the syslog feed, refer to the *SNMP Guide for Cisco Unified ICM/Contact Center Enterprise & Hosted* at:
http://www.cisco.com/en/US/products/sw/custcosw/ps1001/products_installation_and_configuration_guides_list.html

For details on the health monitoring and management capabilities of Unified CCE, review the *Serviceability Best Practices Guide for Cisco Unified ICM/Contact Center Enterprise & Hosted* at:
http://www.cisco.com/en/US/products/sw/custcosw/ps1001/products_installation_and_configuration_guides_list.html

Combining IP Telephony and Unified CCE in the Same Unified CM Cluster

It is possible for a Unified CM cluster to support Cisco Unified IP phones with both normal IP Telephony (office) extensions and Unified CCE (call center) extensions. When running dual-use Unified CM clusters with both IP Telephony and Unified CCE extensions, it is important to ensure all elements of the solution are compatible, as documented in the Cisco Unified Contact Center Enterprise Software Compatibility Guide, available at
http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_device_support_tables_list.html.
It is also important to note that many contact center environments have very stringent maintenance windows. Additionally, Unified CCE agents process far more calls than typical administrator phone users in a Unified CM cluster, so their device weight (or the amount of processing power required per agent) is higher than a typical business phone user. For example, an administrator-only cluster might be able to support 20,000 phones, but a Unified CCE cluster might support only a fraction of those as agents because of the higher call volume and messaging that Unified CM is required to maintain to support those agents. Because of these software and environmental limitations, it might sometimes be advantageous to separate the Unified CM clusters for IP Telephony extensions from the Unified CM clusters for Unified CCE extensions. It is important to consider the environment where Unified CCE is being deployed to determine whether a separate Unified CM cluster is advantageous.

Combining IP Telephony and Unified CCE Extensions on the Same IP Phone

Unified CCE supports only one agent ACD line on the IP phone, which typically will not have voicemail or any call forwarding defined so that Unified CCE can manage and control all calls sent to the agent on this line. Typically, the agent extension is not used as the agent's DID or personal line. A separate line can be assigned to the agent's phone for that purpose, and can be configured with voicemail and other calling features.

The position of the line on the phone determines which line will be answered or used if the agent just picks up the handset. In a typical call center, the ACD line would be the first line on the phone to make it easier for the agent to answer inbound ACD calls and also to ensure that any calls the agent makes using the phone are tracked by the system as external calls for that agent. Additionally, the agent's state will change based upon this line. If the agent picks up the phone to place a call, the agent will be put into not-ready mode and the Unified CCE will not route a call to that agent.

In some cases the agents are knowledge workers, or they do not take as many ACD calls as they do normal extension calls. The call center manager would not want to track all of their phone activity that is not ACD related, and it might be inconvenient for those users to always get the ACD line first when they want to pick up a DID call instead. In this case, the order of the lines might best be reversed, placing the ACD line on the last (or bottom) line appearance on the phone and placing the DID or normal extension on the first line on the phone. This arrangement will allow the users to pick up the phone and answer the first line as well as use this line for all calls they place by default. To answer an ACD call, they will have to select that line on the phone or use the agent desktop to answer that line appearance directly. Also, be aware that they will have to manage their agent state and go into not-ready mode manually when they want to place a call on their normal extension, so that Unified CCE will not attempt to route a call to them while they are on the other line.

It is possible to have a deployment where the agent extension is the same as the agent's DID or personal line. When call waiting is configured on the agent phone, agent-to-agent calls could interrupt a customer call. To prevent this from happening, agent-to-agent routing can be used and the agent-to-agent routing script can be set up to queue or reject the call if the agent is busy. This is a good option if there is a need to see all agent activity and to avoid all interruptions for the agent. The configuration involves using CTI Route Points in Unified CM instead of the agent DID in order to send the calls to Unified CCE for agent-to-agent routing. For ease of configuration and to reduce the number of CTI route points, the Unified CM wildcard feature can be used, although Unified CCE will require distinct routing DNs, one for each agent.

Agent Phones in Countries with Toll-Bypass Regulations

Some countries such as India have telecommunications regulations that require the voice infrastructure to be partitioned logically into two systems: one for Closed User Group (CUG) or Voice over IP (VoIP) to enable communications across the boundaries within the organization, and a second one to access the local PSTN. To ensure adherence of the regulations in such countries, agents typically used to have only one line with access to customer calls only, and they were required to have a different phone (for example, a softphone) to access a VoIP line for contacting fellow teammates or experts located outside the contact center.

The Logical Partitioning feature in Cisco Unified CM provides the same capability through a telephony system to control calls and features on the basis of specific allowed or forbidden configurations. A common telephony system in a contact center environment can provide access to both the PSTN and VoIP networks, therefore configurations are required to provide controlled access and to avoid toll bypass. The Logical Partitioning feature can be enabled and configured in Unified CM to prevent toll-bypass calls, thus allowing agents in a Unified CCE system to use the same phone for receiving customer calls and for making or receiving VoIP calls to and from other people within the organization.

Although this eliminates the need for agents to have a second phone, contact center managers can choose to have a dedicated line or phone for customer calls and can allocate a different line or phone for other calls.

Queuing in a Unified CCE Environment

Call queuing can occur in three distinct scenarios in a contact center:

- New call waiting for handling by initial agent
- Transferred call waiting for handling by a second (or subsequent) agent
- Rerouted call due to ring-no-answer, waiting for handling by an initial or subsequent agent

When planning your Unified CCE deployment, it is important to consider how to handle queuing and requeuing.

Call queuing in a Unified CCE deployment requires use of an IVR platform that supports the SCI interface to Unified CCE. The Unified IP IVR is one such platform. Cisco also offers another IVR platform, Unified CVP, that can be used as a queuing point for Unified CCE deployments. Chapter 2, "Deployment Models" provides considerations for deployments with Unified CVP. Traditional IVRs can also be used in Unified CCE deployments, and Chapter 2, "Deployment Models" also provides considerations for deployments with traditional IVRs.

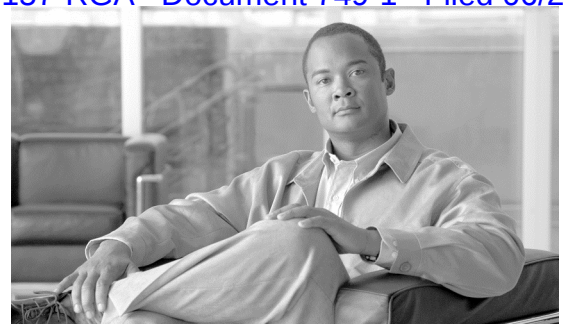
In a Unified CCE environment, an IVR is used to provide voice announcements and queuing treatment while waiting for an agent. The control over the type of queuing treatment for a call is provided by Unified CCE via the SCI interface. The Run VRU Script node in a Unified CCE routing script is the component that causes Unified CCE to instruct the IVR to play a particular queuing treatment.

While the IVR is playing the queuing treatment (announcements) to the caller, Unified CCE waits for an available agent with a particular skill (as defined within the routing script for that call). When an agent with the appropriate skill becomes available, Unified CCE reserves that agent and then instructs the IVR to transfer the voice path to that agent's phone.

Transfers and Conferences in a Unified CCE Environment

Transfers and conferences are commonly used features in contact centers, and they require special attention to ensure the proper system resources are available and configured correctly. (See Chapter 2, "Deployment Models".)

When Call Recording is enabled in the DN configuration for an agent phone, the codec will not be renegotiated when establishing a conference. As a result, if two phones are connected using G.722 and a conference call is initiated, the codec will not be renegotiated to G.711 and a hardware conference bridge or transcoder will be required.



CHAPTER 2

Deployment Models

There are numerous ways that Unified CCE can be deployed, but the deployments can generally be categorized into the following major types or models:

- Single Site
- Multi-Site Centralized Call Processing
- Multi-Site Distributed Call Processing
- Clustering over the WAN

Many variations or combinations of these deployment models are possible. The primary factors that cause variations within these models are as follows:

- Locations of Unified CCE servers
- Locations of voice gateways
- Choice of inter-exchange carrier (IXC) or local exchange carrier (LEC) trunks
- Pre-routing availability
- IVR queuing platform and location
- Transfers
- Traditional ACD, PBX, and IVR integration
- Sizing
- Redundancy

This chapter discusses the impact of these factors (except for sizing) on the selection of a design. With each deployment model, this chapter also lists considerations and risks that must be evaluated using a cost/benefit analysis. Scenarios that best fit a particular deployment model are also noted.

In this chapter, section titles are prefaced by the type of factor discussed in the section. The factors are classified into the following categories:

- IPT — Cisco Unified Communications deployment factors (how Cisco Unified Communications Manager and the voice gateways are deployed)
- Unified CCE — Unified CCE deployment factors (such as what PG is used)
- IVR — IVR and queuing deployment factors (if Unified CVP or Unified IP IVR is used)

A combination of these deployment models is also possible. For example, a multi-site deployment may have some sites that use centralized call processing (probably small sites) and some sites that use distributed call processing (probably larger sites). Examples of scenarios where combinations are likely are identified within each section.

Also in this chapter is a section on integration of traditional ACD and IVR systems into a Unified CCE deployment, with considerations on hybrid PBX/ACD deployments. Sizing and redundancy are discussed in later chapters of this Unified CCE design guide. For more information on the network infrastructure required to support a Unified CCE solution, refer to the latest version of the *Cisco Network Infrastructure Quality of Service Design* guide, available at

http://www.cisco.com/go/designzone

For more information on deployment models for Unified CCE and Cisco Unified Communications, refer to the latest version of the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

http://www.cisco.com/go/ucsrnd

What's New in This Chapter

The following topics are new in this chapter or have changed significantly from previous releases of this document.

- Administration & Data Server deployments.
- Beginning with Unified CCE 8.0(1), the Distributor AW (with or without HDS) is renamed as Administration & Data Server. Multiple Administration & Data Server deployments with different roles are available based on the functionality and amount of reporting data that it can handle.
- Unified CCX 8.0(s) Gateway PG Deployment
- The Unified CCX 8.0(1) moving to the Unified Communications Operating System requires the Unified CCE Gateway PG to be installed on a separate (Windows) server. The deployment model change, for new and existing customers, requires the Unified CCE Gateway PG and the Unified CCX ACMI Manager to be installed on separate boxes.
- Unified CCE 8.0(s) co-located with Unified CCE Gateway PG and Unified CCE System PG
- In Unified CCE 8.0, a new parent/child deployment is supported where the Unified CCE Gateway is installed on the same server as the Unified CCE System PG. In this deployment, the scalability limits for agents and calls are reduced. Refer to Sizing Unified CCE Components and Servers for additional details.
- No 8.0 release for Unified System CCE.

Unified CCE 8.0 supports interoperability with an IPv6-enabled Cisco Unified Communications Manager (Unified CM) cluster.

Note The new and changed information in this chapter is extensive, therefore; read the entire chapter.

General Deployment Options

This section describes options that can apply to many of the specific deployment models listed in the rest of this document. It describes at a high level the trade-offs that can be made when installing the Unified CCE software.

Agent Peripheral Options

There are two types of Agent Peripherals that can be installed to handle Unified CCE agents. This section talks about those two types of peripherals and the strengths and weaknesses of each.

Enterprise Unified CCE Peripheral

In these deployments, the Unified CCE software treats the VRU and Unified CM as separate peripherals. This means that routing must be done once for each peripheral, and Termination Call Detail records are created for each peripheral each time a call touches the peripheral. Translation routes must be used to send calls between the VRU and Unified CM.

The Unified CM PG and VRU PG can be deployed independently, or the Unified CM and VRU can be deployed in a Generic PG with separate PIMs for Unified CM and VRU.

These deployments provide a large degree of flexibility in configuration. For example, this deployment is capable of using either a Unified CVP or a Unified IP IVR attached to the VRU peripheral, and load balancing can be done between multiple IVRs.

These Unified CCE deployments may not be used where Unified CCE is a child to a Unified ICM; a Unified CCE System peripheral deployment must be used for that solution. For more information, see the section on Parent/Child .

Unified CCE System Peripheral

The Unified CCE System peripheral combines the functionality of both the VRU peripherals (up to five Unified IP IVR peripherals) and a single Unified CM peripheral together into a single logical Unified ICM peripheral. The Unified CCE treats these Unified IP IVR and Unified CM peripherals as a single peripheral, eliminating the need to translation-route calls to the Unified IP IVR for treatment and queueing. If multiple Unified IP IVRs are configured, the Unified CCE System peripheral automatically load-balances calls between the Unified IP IVRs that have available capacity.

Additionally, because the Unified CCE System PG is a single peripheral, Termination Call Detail (TCD) records and other reporting data include the information for the call during the entire time it is on the peripheral. Instead of getting up to three TCDs for each call (one for the original route, one for the IVR, and one for the agent handle time), only a single record is created with the Unified CCE System PG.

The Unified CCE System PG does not support Unified CVP, therefore all queuing and treatment in the Unified CCE System PG is done using Unified IP IVR. Note that a separate Unified CVP on its own PG can be used in conjunction with the Unified CCE System Peripheral.

Unified CCE: Administration & Data Server

Beginning with Unified CCE 8.0(1), the Distributor AW (with or without HDS) is renamed as Administration & Data Server. Multiple Administration & Data Server deployments with different roles are available, based on the functionality and amount of reporting data that it can handle.

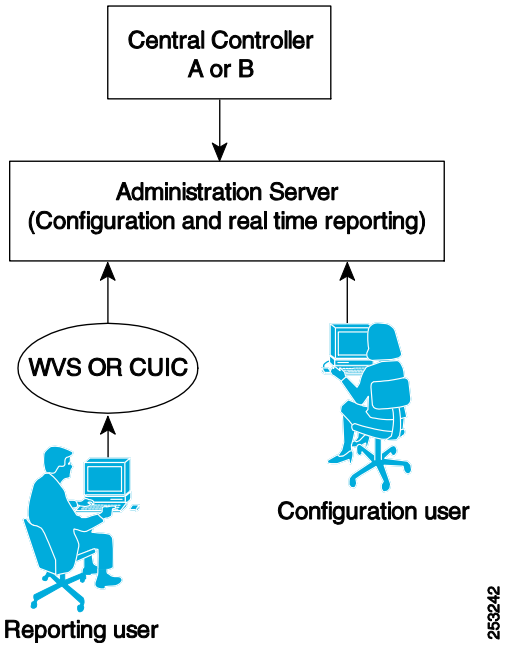
Roles:

The Administration & Data Servers are classified into the following roles based on the system configuration and the call load that it can handle:

Administration Server (Configuration and Real-Time Reporting)

This role is similar to the former "Distributor AW" model, which provides the ability for configuration changes as well as real-time reporting. The real-time reporting is supported using either WebView server or Cisco Unified Intelligent Center (Reporting client). (See Figure 2-1.) No historical reporting is supported.

Figure 2-1 Configuration and Real-Time Reporting

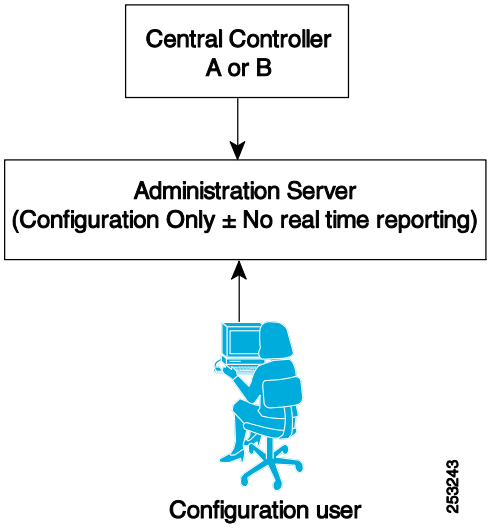


Administration Server (Configuration-Only AW)

In this Administration & Data Server deployment role, the HDS is not enabled and real-time reporting is turned OFF. This distributor deployment provides the ability for configuration changes only. No real-time and historical reporting is supported. (See Figure 2-2.)

This deployment role allows Contact Center Hosted using CCMP to configure a specific Unified CCE Customer Instance through the ConAPI interface. The load is low enough on such a lightweight Administration & Data Server that a single server is sufficient if deployed using VMware. For historical and real-time data, the customer can use a shared AW-HDS or AW-HDS-DDS.

Figure 2-2 Configuration-Only AW

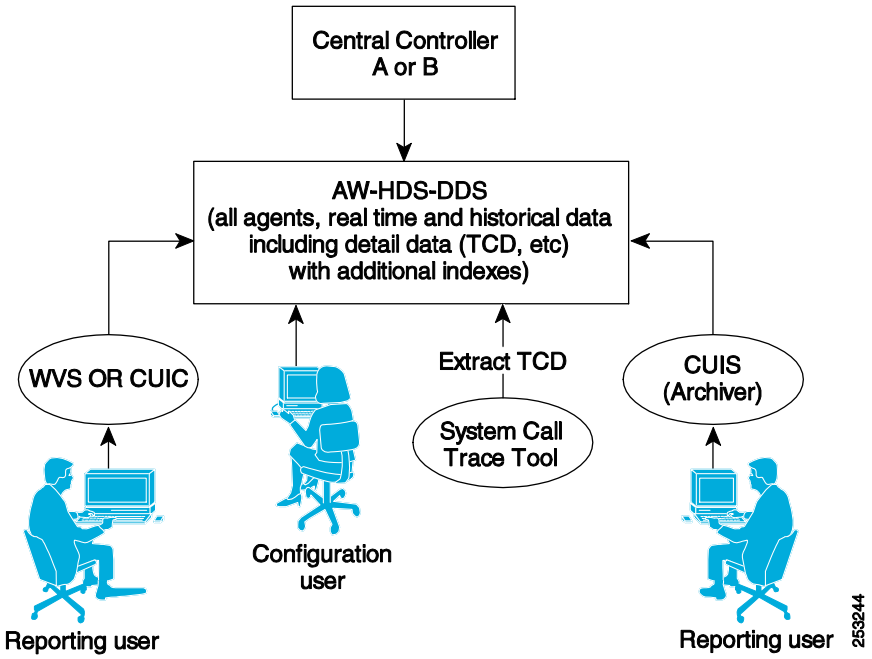


Administration Server, Historical Data Server, and Detail Data Server (AW-HDS-DDS)

This Administration & Data Server deployment role is similar to the existing "Distributor AW with HDS" model, which provides the ability for configuration changes as well as both real-time and historical reporting. The real-time and historical reporting is supported using either a WebView server or Cisco Unified Intelligent Center (CUIC Reporting client). The call detail and call variable data are supported for custom reporting data extraction to meet the requirements for the System Call Trace tool and to feed historical data. (See Figure 2-3.)

Note Cisco Unified Intelligent Center (CUIC) is not part of the out-of-the-box solution.

Figure 2-3 Administration Server, Historical Data Server, and Detail Data Server (AW-HDS-DDS)



Administration Server and Historical Data Server (AW-HDS)

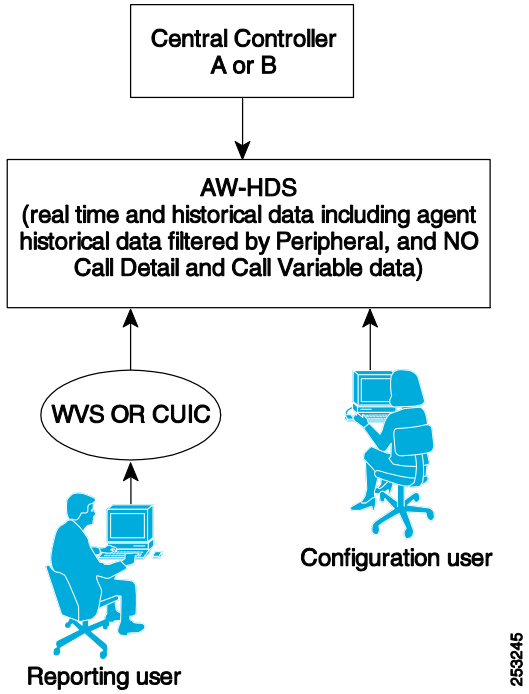
This Administration & Data Server deployment role provides the ability for configuration changes as well as both real-time and historical reporting. The real-time and historical reporting is supported using either a WebView server or Cisco Unified Intelligent Center (Reporting client). (See Figure 2-4.)

Note Cisco Unified Intelligent Center (CUIC) is not part of the out-of-the-box solution.

In addition, the following features are disabled and not supported:

- The Call Detail, Call Variable and Agent State Trace data.
- Custom reporting data extraction procedure.
- Data extraction for System Call Trace tool.
- Feed to CUIS (Archiver)

Figure 2-4 Administration Server and Historical Data Server (AW-HDS)



Historical Data Server and Detail Data Server (HDS-DDS)

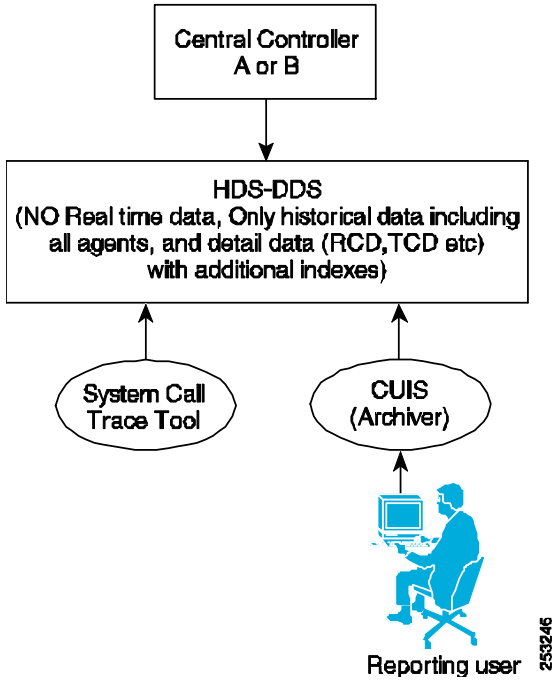
This Administration & Data Server deployment role provides support mainly for historical reporting, Call Detail data extraction for the System Call Trace tool. This deployment also includes configuration data available for historical reporting. (See Figure 2-5.)

In addition, the following features are disabled and not supported:

- Real-time data reporting.
- Ability to make configuration changes.

This deployment role is limited to one per Logger side.

Figure 2-5 Historical Data Server and Detail Data Server (HDS-DDS)



Deployment Options

The following deployment options are based on the varying capacity in terms of number of agents, call load, number of reporting users, and other operating conditions listed in the Sizing Unified CCE Components and Servers section 10-3:

Small to Medium:

This deployment option is applicable to the following limitations:

- Sizing within the operating conditions listed in the Sizing Unified CCE Components and Servers section 10-3.
- A maximum of 4000 active agents.
- A maximum of 400 reporting users.

The number of HDS is limited to 2 AW-HDS-DDS per Logger side.

Large:

This deployment option is applicable to the following limitations:

- Sizing within the operating conditions listed in the Sizing Unified CCE Components and Servers section 10-3.
- More than 4000 active agents.
- More than 400 reporting users.

The number of Administration & Data Server roles is limited to 3 AW-HDS and 1 HDS-DDS or AW-HDS-DDS per Logger side.

Both WebView Server and Cisco Unified Intelligent Center (Reporting client) are supported; however Cisco Unified Intelligent Center is not part of the out-of-the-box solution.

Unified System CCE

There is no 8.0 release for Unified System CCE; it will remain at 7.5. Users can migrate their Unified System CCE 7.1, 7.2, or 7.5 system directly to Unified CCE 8.0 using the Unified CCE installer. Unified CCE 8.0 supports Parent/Child integrations to Unified System CCE 7.x.

Parent/Child

The Unified CCE Gateway PG allows Unified CCE or Unified CCX to appear as a traditional ACD connected to the Unified ICM system. The Unified CCE Gateway PG does this by providing a PG to the Unified ICM system that communicates to the CTI interface of Unified CCE System PG or Unified CCX.

When the Unified CCE Gateway PG is used in a deployment, its relationship of Unified ICM is termed a *parent* and Unified CCE is called the *child*:

- Parent

  The Unified ICM system that serves as the network or enterprise routing point. The child looks like an ACD to the parent, which uses the appropriate Unified CCE Gateway PG (Enterprise or Express) to communicate to the CTI interface on the child Unified CCE. The parent can perform all functions that a Unified ICM can usually perform, including pre- and post-routing and end-to-end call tracking using translation routes.

- Child

  The Unified CCE System PG or Unified CCX system that is set up to function as an ACD. The child can receive calls that are translation-routed from the parent, but it is not aware of any other peripherals attached to the parent. The child can also post-route calls from the Unified CCE to the parent, where the call can be handled like any other Unified ICM call. For example, the call could be translation-routed to any (TDM or IP) ACD controlled by the Unified ICM or queued in the Unified ICM network queue point with Unified CVP.

In the parent/child model, the child Unified CCE is configured to function completely on its own and does not need the connection to the parent to route calls to agents. This independence provides complete local survivability for mission-critical contact centers if the network between the child and parent goes down or if there is a problem with the parent or the Unified CCE Gateway PG connection.

Configuration objects entered into the child system can automatically be sent to the parent Unified ICM and inserted into the Unified ICM configuration, thus eliminating the need to configure objects twice, once in the local ACD and again to match the configuration in the Unified ICM itself for routing and reporting. This functionality can also be turned off for situations where the customer does not want automatic configuration updates, such as with an outsourcer using the Unified CCE child system where not all of the agents, skill groups, and call types on that child system apply to the customer's Unified ICM system.

The Unified CCE Gateway PG can connect to a Unified CCE child that is using the Unified CCE System PG or to Unified CCX. If the Unified CCE child has multiple Unified CCE System PGs and peripherals, a separate Unified CCE Gateway PG peripheral must be installed and configured for each one in the Unified ICM parent system. When deployed on a separate server, a Unified CCE Gateway PG can manage multiple child Unified CCE peripherals or multiple child Unified CCX systems, with up to five child systems. Note, the Unified CCE Gateway PG does not support Unified CCE System PG and Unified CCX integration on the same CCE Gateway PG instance.

In the Unified CCE child, either IP IVR or Unified CVP may be deployed for call treatment and queuing. If Unified CVP is deployed, an additional VRU PG must be configured, and this model does not follow the single peripheral model used when IP IVR is deployed. For this reason, information on calls queued at the child (and queue time of a call) is not available on the parent, so any computation involving queue time will be inaccurate (for example, minimum expected delay (MED) and average answer wait time).

Special Note on Network Consultative Transfer (NCT) for Parent/Child Systems:

One restriction of parent/child is that calls terminating on child systems cannot be transferred via network consultative transfer (NCT) through the NIC or any other routing client on the parent. Although NCT works for TDM ACDs, and at first glance parent/child seems virtually identical in architecture, parent/child is not the same. For a TDM PG, the CTI-Server is connected to the PG ACD, which is part of the parent system. This would be the equivalent of having a CTI-Server connected to the Gateway PG. To think of it another way, it is like using CTI directly to an ACD instead of the CTI Server, in which case network consultative transfer is not possible either. In parent/child deployments, CTI is connected to the child PG. Having CTI connected to the child PG does not provide the necessary network call ID and other information necessary to allow network consultative transfer.

Note, however, that network blind transfer is still possible using any client (for example, Unified CVP or a NIC) on the parent system when a post route is initiated to the parent system from the child.

SIP Support

Unified IP IVR is notified of caller entered digits (DTMF input) by way of JTAPI messages from Unified CM. Unified IP IVR and Unified QM do not support mechanisms to detect in-band DTMF digits. In deployments with SIP voice gateways or SIP phones that support only in-band DTMF (or are configured to use in-band DTMF per RFC 2833), Unified CM must invoke an MTP resource to convert the in-band DTMF signaling to out-of band signaling so that the Unified IP IVR can be notified of the caller entered digits. Therefore, in environments that include SIP phones or gateways, it is necessary to provision sufficient MTP resources. Keep this in mind if the phones need to interact with Unified IP IVR. Likewise, CTI ports do not support in-band DTMF (RFC 2833). The Mobile Agent feature relies on CTI ports, so MTP resources are required when in-band DTMF (RFC 2833) is negotiated.

SIP trunking is supported using CVP deployments with Cisco IOS gateways (IOS GW) and Unified Border Element (CUBE) (CUBE). For more information on deployment models for Unified CCE and Cisco Unified Communications, refer to the latest version of the Cisco Unified Communications Solution Reference Network Design (SRND) guide, available at

http://www.cisco.com/go/ucsrnd

Q.SIG Support

Cisco Unified CCE does not support using Q.SIG trunks with the Unified CM deployment.

IPv6 Support

Unified CCE 8.0 supports interoperability with an IPv6-enabled Unified CM cluster. All the Unified CCE components run with IPv4 enabled, which includes IP-IVR, Unified CVP, and the CTIOS Agent Desktops, and Agent Phones. The Unified CCE Agent PG integration with the Unified CM CTI Manager uses IPv4.

Caller phones or voice gateways could run IPv4 or IPv6. If the caller's environment is IPv6 only, then Media Termination Point (MTP) resources are required for call treatment using IP-IVR and Unified CVP VXML Gateways.

Agent phones must run IPv4; IPv6 is not supported for agent phones. If the caller's phone is IPv6, then an MTP must also be inserted between the resources.

Mobile Agent and Outbound Option endpoints (CTI ports and Dialer ports) must be configured as IPv4 devices.

Service Advertisement Framework Call Control Discovery (SAF CCD)

The Service Advertisement Framework Call Control Discovery feature is a way to replace the need for gatekeepers and SIP proxies. The JTAPI impact is a new reason code when calls are redirected from another cluster back to the local PSTN. For more details, refer to the *Cisco Unified Communications SRND*, available at

http://www.cisco.com/go/ucsrnd

Cisco Unified Mobile Agent

For deployments using Cisco Unified Mobile Agent, it is important to consider the location of the voice gateways that will be used to call agents because their location has design considerations for silent monitoring, call admission control, and other areas. For design guidance and considerations for implementing Cisco Unified Mobile Agent, see the chapter on Cisco Unified Mobile Agent Architecture , page 6-1.

CTI-OS Multi-Server Support

Cisco Unified CCE supports multiple CTI OS servers connecting to a single CTI Server. Up to ten CTI OS servers are allowed per PG.

This deployment provides the following benefits:

- Simplification: multiple CTI OS Servers can be configured to use the same CTI Server.
- Many small sites may use a single PG with multiple PIMs rather than each requiring its own PG.
- Reduced box count because all of the PG processes, including the PIM and CTI OS Server processes, are running on the same box.
- Increased Scalability of a single Unified CCE PG because multiple PIMs under a single PG are used to connect to different Unified CM clusters.

This deployment has the following requirements:

CTI OS Servers must reside on the same server as the rest of the PG processes.

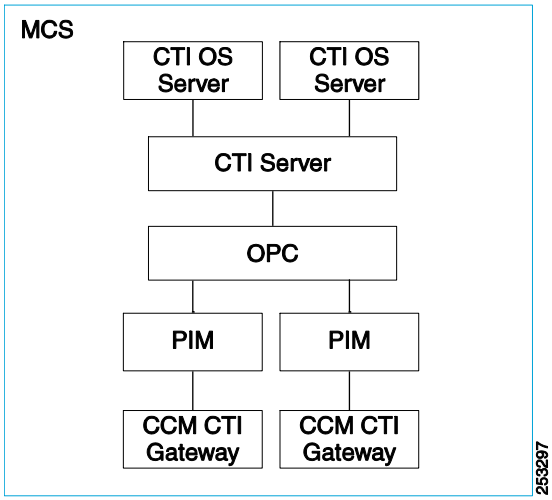
- Each PG can be configured for only one peripheral type.
- ARS and ERS peripheral types are supported in this deployment model.
- Multi-instance deployments cannot add more than one CTI OS Server per instance.
- This deployment model may be used with Unified CCE peripherals.

Note PGs are typically co-located with a peripheral. Allowing multiple peripherals per PG could result in some peripherals being situated remotely from the PG. This is not supported for some peripherals and remains unsupported in this case. For example, the Unified CM (Enterprise and System deployments) would not be aggregated in a single PG unless all the ACD and Unified CM peripherals were co-located on a local LAN with the PG. In general, the deployment rules associated with ACD integrations on a PG still apply. For those deployments supporting remote PGs, all network requirements (including bandwidth, latency, and availability requirements) must be met.

When this deployment model is implemented, scaling is reduced to 75% of the scaling capacity for a single CTI OS. For example, if a given configuration supports 1000 agents with a single CTI OS server, it will support 750 agents using multiple CTI OS servers. This is due to the extra overhead of extra CTI OS processes and to the extra processing load incurred by the CTI Server due to the extra clients. The

exception to this is when this feature is used for supporting over 2000 agents (the CTI Manager limit) on Unified CCE. (See Figure 2-6 for an example.) Note that this deployment is supported only when using the Unified CCE PG, and it does not support a VRU under the same PG (no Generic type supported).

Figure 2-6 Multiple CTI OS Servers



CAD Multi-Server Support

Multiple instances of CAD services are supported on a single PG. The CAD services must reside on separate servers, but they do not require a separate PG for each one. Up to 10 instances of CAD services are allowed per PG.

The benefit of this type of deployment is that the additional instances of CAD services required to achieve a specific agent number do not require a separate PG. While the separate instances of CAD services do have to be deployed on separate servers, the hardware requirements for CAD services are minimal. For information on those requirements, refer to the *Cisco CAD Installation Guide/Cisco Unified Contact Center Enterprise and Hosted Release 8.0*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html

Virtualization Support

For detailed information on Unified Communications applications running in virtual machines on UCS hardware, please see the Unified Communications Virtualization Doc Wiki page at:

http://docwiki.cisco.com/wiki/Unified_Communications_Virtualization

IPT: Single Site

A single-site deployment refers to any scenario where all voice gateways, agents, desktops, phones, and call processing servers (Unified CM, Unified ICM, Unified CCE, and Unified IP IVR or Cisco Unified Customer Voice Portal (Unified CVP)) are located at the same site and have no WAN connectivity between any Unified CCE software modules. Figure 2-7 illustrates this type of deployment using the Unified CCE model.

Figure 2-7 Single-Site Deployment

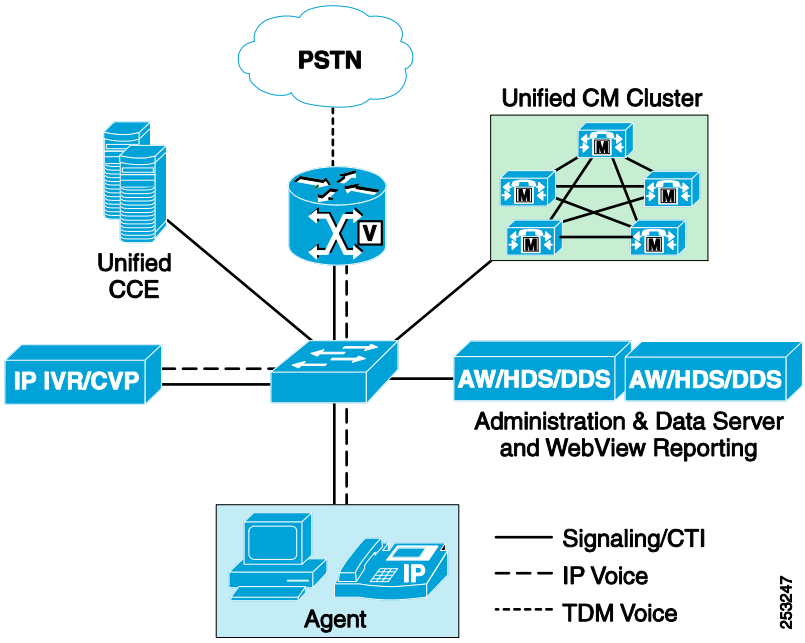


Figure 2-7 shows a Unified IP IVR, a Unified CM cluster, redundant Unified CCE servers, two Administration & Data Servers with WebView, and a direct connection to the PSTN from the voice gateways. The Unified CCE server in this scenario is running the following major software processes:

- Call Router
- Logger and Database Server
- Unified CCE System PG with Unified CM Peripheral Interface Manager (PIM) and Unified IP IVR PIM
- CTI Server
- CTI Object Server (CTI OS)
- Optionally, Cisco Agent Desktop (CAD) servers could be co-located on the Unified CCE servers as well.

Optionally the Central Controller and Unified CCE System PG (and so forth) can be split onto separate servers. For information on when to install the Central Controller and PG on separate servers, refer to Chapter 10, "Sizing Unified CCE Components and Servers".

Traditional Unified CCE must be deployed in a redundant fashion. Simplex deployments are supported only for lab or non-production deployments. For information on Unified CCE redundancy, refer to Chapter 3, "Design Considerations for High Availability".

The number of Unified CM nodes and the hardware model used is not specified along with the number of Unified IP IVRs. For information on determining the number and type of servers required, refer to Chapter 10, "Sizing Unified CCE Components and Servers".

Also not specified in this model is the specific data switching infrastructure required for the LAN, the type of voice gateways, or the number of voice gateways and trunks. Cisco campus design guides and Cisco Unified Communications design guides are available to assist in the design of these components. Chapter 9, "Sizing Contact Center Resources " discusses how to determine the number of gateway ports.

Another variation in this model is to have the voice gateways connected to the line side of a PBX instead of the PSTN. Connection to multiple PSTNs and a PBX all from the same single-site deployment is also possible. For example, a deployment can have trunks from a local PSTN, a toll-free PSTN, and a traditional PBX/ACD. For more information, see Traditional ACD Integration , page 2-49, and Traditional IVR Integration , page 2-52.

This deployment model also does not specify the type of signaling (ISDN, MF, R1, and so on) to be used between the PSTN and voice gateway, or the specific signaling (H.323, SIP or MGCP) to be used between the voice gateway and Unified CM.

The amount of digital signal processor (DSP) resources required for placing calls on hold, consultative transfers, and conferencing is also not specified in this model. For information on sizing of these resources, refer to the latest version of the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

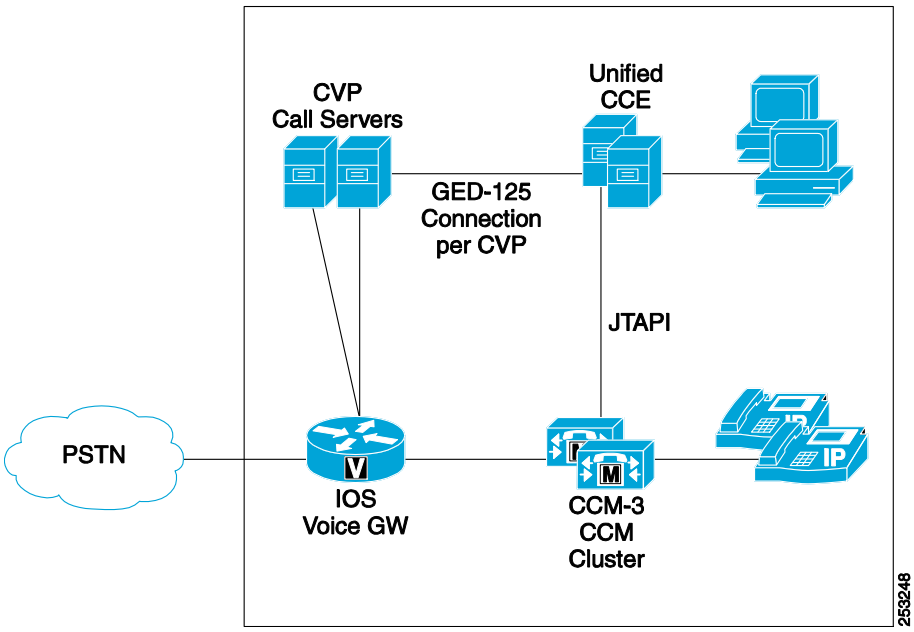
http://www.cisco.com/go/ucsrnd

The main advantage of the single-site deployment model is that there is no WAN connectivity required. Given that there is no WAN in this deployment model, there is generally no need to use G.729 or any other compressed Real-Time Transport Protocol (RTP) stream, so transcoding would not be required.

Unified CCE: Unified CCE System PG

In this deployment model, the agent PG that is deployed is a Unified CCE System PG. Only a single peripheral is needed to handle both the Unified CM and any Unified IP IVRs that may exist. This peripheral unifies the appearances of the multiple PIMs and also handles the load balancing of calls between multiple Unified IP IVRs. Alternatively, this model may be configured to use Unified CVP. When Unified CVP is used, its connectivity to Cisco Unified Presence handles load balancing by distributing the incoming calls among Unified CVP Call Servers. In this deployment, the VRU PIMs (up to 10) communicating with the Unified CVP Call Server(s) reside on their own PG and not under the Unified CCE System PG. Figure 2-8 shows a single-site deployment utilizing Unified CVP instead of IP-IVR in a Unified CCE system. In this model, no longer do all calls reside under a single peripheral; Unified CVP is its own peripheral(s).

Figure 2-8 Single-Site Deployment Using Unified CVP



When using this configuration, the VRU PGs must be deployed redundantly with one to ten Unified CVP Call Servers, depending upon the number of agents and call volume required. From a practical point of view, this deployment is similar to a single-PG Unified CCE system with the major exception of greatly simplified setup and configuration.

IVR: Treatment and Queuing with Unified IP IVR

In this deployment model, all initial and subsequent queuing is done on the Unified IP IVR. Up to five Unified IP IVRs can be deployed in this model (with the Unified CCE System PG). The Unified IP IVRs are placed behind Unified CM, using Unified CM's dial plan and call switching under control of Unified CCE. All calls come into a CTI Route Point on Unified CM, controlled by Unified CCE, and are then automatically translation-routed to the Unified IP IVR by the Unified CCE System PG. The Unified CCE handles load balancing between available Unified IP IVR ports, and configuring translation routes between the Unified IP IVR and Unified CM is not needed.

IVR: Treatment and Queuing with Unified CVP

Unified CVP could be used to provide the call treatment and queuing in this model as well. Because Unified CVP is not part of the Unified CCE System PG peripheral, translation routes must be configured to transfer calls with call data between the peripherals.

In this deployment model, all initial and subsequent queuing is done using Unified CVP. A single server may be used, with all Unified CVP processes co-located on that server. Multiple servers, on the other hand, allow scaling and redundancy. For more information about redundancy, see Chapter 3, “Design Considerations for High Availability”.

For more information about Unified CVP, refer to the *Cisco Unified Customer Voice Portal Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Unified CCE: Enterprise Unified CCE PG

In these deployment models the Enterprise Unified CCE peripheral is used to handle interactions with Unified CM, and a separately configured VRU peripheral is used to handle interactions with the Unified IP IVR or Unified CVP.

IVR: Treatment and Queuing with Unified IP IVR

In this deployment model, all initial and subsequent queuing is done on the Unified IP IVR. If multiple Unified IP IVRs are deployed, use Unified CCE to load-balance calls across those Unified IP IVRs. Translation routes must be configured manually between the Unified CM peripheral and the Unified IP IVR peripheral(s) and used to move calls and data between Unified CM and the Unified IP IVRs. Load balancing is done manually in the Translation Route To VRU node in the Unified CCE call routing script.

IVR: Treatment and Queuing with Unified CVP

Unified CVP could be used to provide the call treatment and queuing in this model as well. Unified CVP would have its own VRU PG, either loaded on the same server as the Unified CM PG or part of a Generic PG combination.

In this deployment model, all initial and subsequent queuing is done using Unified CVP. A single server may be used, with all Unified CVP processes co-located on that server. Multiple servers, on the other hand, allow scaling and redundancy. For more information about redundancy, see the Chapter 3, "Design Considerations for High Availability".

For more information about Unified CVP, refer to the *Cisco Unified Customer Voice Portal Solution Reference Network Design (SRND)*, available at

http://www.cisco.com/go/ucsrnd

Unified CCE: Transfers

In this deployment model (as well as in the multi-site centralized call processing model), both the transferring agent and target agent are on the same peripheral. This also implies that both the routing client and the peripheral target are the same peripheral. The transferring agent generates a transfer to a particular dialed number configured as a CTI Route Point in Unified CM (for example, looking for any specialist in the specialist skill group).

The Agent peripheral (either the Unified CCE System peripheral or the Enterprise Unified CCE peripheral) will generate a route request to the Call Router. The Call Router will match the dialed number to a call type and activate the appropriate routing script. The routing script looks for an available specialist.

If a target agent (specialist) is available to receive the transferred call, then the Call Router will return the appropriate label, based on Agent Target Rules (Dynamic) or a DeviceTarget Label (Static), to the requesting routing client (the Agent peripheral). In this scenario, the label is typically just the extension of the phone where the target agent is currently logged in. Upon receiving the route response (label), the Unified CM PIM will then initiate the transfer by sending a JTAPI transfer request to Unified CM.

At the same time that the label is returned to the routing client, pre-call data (which includes any call data that has been collected for this call) is delivered to the peripheral target. In this scenario, the routing client and peripheral target are the same Agent peripheral. This is because the transferring agent and the target agent are both associated with the same peripheral. In some of the more complex scenarios to be discussed in later sections, sometimes the routing client and peripheral target are not the same.

If a target agent is not available to receive the transferred call, then the Call Routing script is typically configured to transfer the call to an IVR so that queue treatment can be provided. In this scenario the logic in the Unified CCE System PG differs from the logic in the Unified CCE PG if the IP-IVR variant is used.

In both cases the label is a dialed number that will instruct Unified CM to transfer the call to an IVR. The translation-route or correlationID is not needed when using the Unified CCE System peripheral but is needed when deploying Unified CVP.

IPT: Multi-Site with Centralized Call Processing

A multi-site deployment with centralized call processing refers to any scenario where call processing servers (Unified CM, Unified CCE, and Unified IP IVR or Unified CVP) are located at the same site, while any combination of voice gateways, agents, desktops, and phones are located remotely across a WAN link or centrally. Figure 2-9 illustrates this type of deployment.

There are two variations of this IPT model:

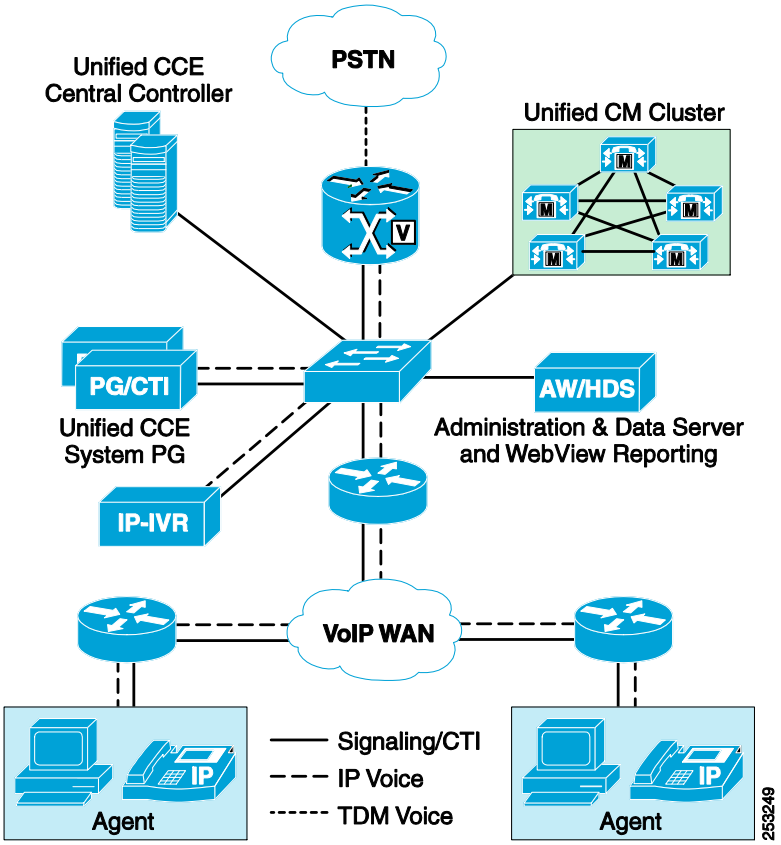
- IPT: Centralized Voice Gateways
- IPT: Distributed Voice Gateways

IPT: Centralized Voice Gateways

If an enterprise has small remote sites or offices in a metropolitan area where it is not efficient to place call processing servers or voice gateways, then this model is most appropriate. As sites become larger or more geographically dispersed, use of distributed voice gateways might be a better option.

Figure 2-9 illustrates this model using a Unified CCE deployment. The illustration shows the deployment using IP-IVR, but it could also use Unified CVP instead of IP-IVR.

Figure 2-9 Multi-Site Deployment with Centralized Call Processing and Centralized Voice Gateways



Advantages

- Only a small data switch and router, IP phones, and agent desktops are needed at remote sites where only a few agents exist, and only limited system and network management skills are required at remote sites.
- No PSTN trunks are required directly into these small remote sites and offices, except for local POTS lines for emergency services (911) in the event of a loss of the WAN link.
- PSTN trunks are used more efficiently because the trunks for small remote sites are aggregated.
- Unified CCE Queue Points (Unified IP IVR or Unified CVP) are used more efficiently because all Queue Points are aggregated.
- No VoIP WAN bandwidth is used while calls are queuing (initially or subsequently). Calls are extended over the WAN only when there is an agent available for the caller.

As with the single-site deployment model, all the same options exist when using Unified CCE configurations. For example, multi-site deployments can run the Unified CCE software all on the same server or on multiple servers. The Unified CCE software can be deployed either with the Unified CCE System PG or the Unified CCE PG. The number of Unified CM and Unified IP IVR or Unified CVP servers is not specified by the deployment model, nor are the LAN/WAN infrastructure, voice gateways, or PSTN connectivity. For other variations, see IPT: Single Site, page 2-11.

Best Practices

- VoIP WAN connectivity is required for RTP traffic to agent phones at remote sites.
- RTP traffic to agent phones at remote sites might require compression to reduce VoIP WAN bandwidth usage. It may be desirable for calls within a site to be uncompressed, so transcoding might also be required depending upon how the Cisco Unified Communications deployment is designed.
- Skinny Client Control Protocol (SCCP) or SIP call control traffic from IP phones to the Unified CM cluster flows over the WAN.
- CTI data to and from the Unified CCE Agent Desktop flows over the WAN. Adequate bandwidth and QoS provisioning are critical for these links.
- Because there are no voice gateways at the remote sites, customers might be required to dial a long-distance number to reach what would normally be a local PSTN phone call if voice gateways with trunks were present at the remote site. This situation could be mitigated if the business requirements are to dial 1-800 numbers at the central site. An alternative is to offer customers a toll-free number to dial, and have those calls all routed to the centralized voice gateway location. However, this requires the call center to incur toll-free charges that could be avoided if customers had a local PSTN number to dial.
- The lack of local voice gateways with local PSTN trunks can also impact access to 911 emergency services, and this must be managed via the Unified CM dial plan. In most cases, local trunks are configured to dial out locally and for 911 emergency calls.
- Unified CM locations-based call admission control failure will result in a routed call being disconnected. Therefore, it is important to provision adequate bandwidth to the remote sites. Also, an appropriately designed QoS WAN is critical.
- NOTE: For calls controlled by Unified CVP, call admission control failures can be recovered with the proper Unified CCE Scripting configuration with the Unified CCE Router Requery feature enabled.
- Automated Alternate Routing (AAR) provides a mechanism to reroute calls through the PSTN or other network by using an alternate number when Unified CM blocks a call due to insufficient location bandwidth. For deployments with Unified CVP ingress call control, do not enable AAR. This allows Unified CVP Router Requery to take control of the call in the event of a failure or timeout. For deployments using IP-IVR, enable AAR to route the call through the PSTN or other network component using an alternative number.

IVR: Treatment and Queuing with Unified IP IVR

As in the single-site deployment, all call queuing is done on the Unified IP IVR at a single central site. While calls are queuing, no RTP traffic flows over the WAN. If requeuing is required during a transfer or reroute on ring-no-answer, the RTP traffic flow during the queue treatment also does not flow over the WAN. This reduces the amount of WAN bandwidth required to the remote sites.

IVR: Treatment and Queuing with Unified CVP

In this model, Unified CVP is used in the same way as Unified IP IVR.

Unified CCE: Transfers

Transfers in this scenario are, from the point of view of the contact center, the same as in the single-site scenario. Therefore, the same call and message flows will occur as in the single-site model, whether the transferring agent is on the same LAN as the target or on a different LAN. The only differences are that QoS must be enabled and that appropriate LAN/WAN routing must be established. For details on provisioning your WAN with QoS, refer to the latest version of the *Cisco Network Infrastructure Quality of Service Design* guide, available at

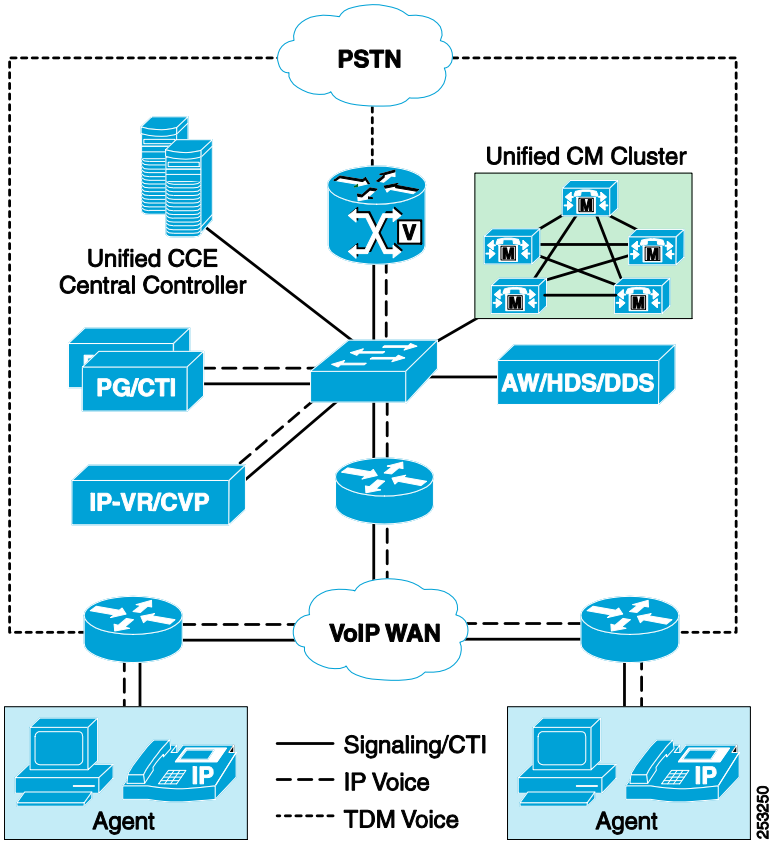
http://www.cisco.com/go/designzone

During consultative transfers where the agent (not the caller) is routed to a Unified IP IVR port for queuing treatment, transcoding is required because the Unified IP IVR can generate only G.711 media streams.

IPT: Distributed Voice Gateways

A variation of the centralized call processing model can include multiple ingress voice gateway locations. This distributed voice gateway model might be appropriate for a company with many small sites, each requiring local PSTN trunks for incoming calls. This model provides local PSTN connectivity for local calling and access to local emergency services. Figure 2-10 illustrates this model.

Figure 2-10 Multi-Site Deployment with Centralized Call Processing and Distributed Voice Gateways



In this deployment model, shown with Unified IP IVR for queuing and treatment, it might be desirable to restrict calls arriving at a site to be handled by an agent within that site, but this is not required. By restricting calls to the site where it arrived, the following conditions apply:

- VoIP WAN bandwidth is reduced for calls going to agents from the ingress voice gateway.
- Calls will still cross the VoIP WAN during the time they are in queue or are receiving treatment from the centralized Unified IP IVRs.
- Customer service levels for calls arriving into that site might suffer due to longer queue times and handling times.
- Longer queue times can occur because, even though an agent at another site is available, the Unified CCE configuration may continue to queue for an agent at the local site only.
- Longer handling times can occur because, even though a more qualified agent exists at another site, the call may be routed to a local agent to reduce WAN bandwidth usage.

In order to restrict a call to the site at which it arrived in this deployment model, it is necessary to create separate skill groups for agents at each location. In order to route a call to any agent in a given skill regardless of location, the location-specific skill groups can be combined using an enterprise skill group.

It is important for deployment teams to carefully assess the trade-offs between operational costs and customer satisfaction levels to establish the right balance on a customer-by-customer basis. For example, it may be desirable to route a specific high-profile customer to an agent at another site to reduce their queue time and allow the call to be handled by a more experienced representative, while another customer may be restricted to an agent within the site where the call arrived.

A Unified CCE deployment may actually use a combination of centralized and distributed voice gateways. The centralized voice gateways can be connected to one PSTN carrier providing toll-free services, while the distributed voice gateways can be connected to another PSTN carrier providing local phone services.

Inbound calls from the local PSTN could be both direct inward dial (DID) and contact center calls. It is important to understand the requirements for all inbound and outbound calling to determine the most efficient location for voice gateways. Identify who is calling, why they are calling, where they are calling from, and how they are calling.

CVP Call Treatment and Call Routing

In multi-site environments with distributed voice gateways, Unified CVP can be used to leverage the ingress voice gateways at the remote sites as part of the Unified CCE system. Unified CVP treats and queues calls locally in the ingress voice gateway rather than requiring the call to cross the VoIP WAN to a centralized queue platform. Unified CVP provides call treatment (VRU) using the VoiceXML Browser built into the Cisco IOS voice gateway. Only call signaling passes over the WAN to instruct the remote site voice gateway how to treat, queue, and transfer the call to an agent. In these models, pre-routing to the site might not be necessary because Unified ICM/CCE takes control of the call as soon as it arrives at the site. Basic carrier percent allocation can be used to allocate calls to the sites, and failover (rollover) trunks can be used to address local failures as needed.

Traditional Prerouting

In a traditional deployment with Unified ICM (parent/child or hybrid) with multi-site deployments and distributed voice gateways, the Unified ICM pre-routing capability can also be used to load-balance calls dynamically across the multiple sites. For a list of PSTN carriers that offer Unified ICM pre-routing services, refer to the *Pre-installation Planning Guide for Cisco ICM Enterprise & Hosted Editions*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html

In multi-site environments where the voice gateways have both local PSTN trunks and separate toll-free trunks delivering contact center calls, the Unified ICM pre-routing software can load-balance the toll-free contact center calls around the local contact center calls. For example, suppose you have a two-site deployment where Site 1 currently has all agents busy and many calls in queue from locally originated calls, and Site 2 has only a few calls in queue or maybe even a few agents currently available. In that scenario, you could have Unified ICM instruct the toll-free provider to route most or all of the toll-free calls to Site 2. This type of multi-site load balancing provided by Unified ICM is dynamic and automatically adjusts as call volumes change at all sites.

Just as in the two previous deployment models, much variation exists in the number and types of Unified ICM/CCE, Unified CM, and Unified IP IVR or Unified CVP servers; LAN/WAN infrastructure; voice gateways; PSTN connectivity; and so forth.

Advantages of Using Distributed Voice Gateways

- Only limited systems management skills are needed for the remote sites because most servers, equipment, and system configurations are managed from a centralized location.
- The Unified CVP or Unified ICM/CCE pre-routing can be used to load-balance calls across sites, including sites with local PSTN trunks in addition to toll-free PSTN trunks.
- No WAN RTP traffic is required for calls arriving at each remote site that are handled by agents at that remote site.
- Unified CVP provides call treatment and queueing at the remote site using the VoiceXML Browser in Cisco IOS on the voice gateway itself, thus eliminating the need to move the call over the VoIP WAN to a central queue and treatment point.

Best Practices

- The Unified IP IVR or Unified CVP, Unified CM, and PGs (for both Unified CM and IVR or Unified CVP) are co-located. In this model, the only Unified CCE communications that can be separated across a WAN are the following:
  - Unified Central Controller to PG
  - PG to Unified CCE Agent Desktops
  - Unified CM to voice gateways
  - Unified CM to phones
  - Unified CVP Call Control Server to remote voice gateway (call control)
- If calls are not going to be restricted to the site where calls arrive, or if calls will be made between sites, more RTP traffic will flow across the WAN. It is important to determine the maximum number of calls that will flow between sites or locations. Unified CM locations-based call admission control failure will result in a routed call being disconnected (rerouting within Unified CM is not currently supported). Therefore, it is important to provision adequate bandwidth to the remote sites, and appropriately designed QoS for the WAN is critical. Calls that are treated by IP IVR at the central site must also be considered.

Note For calls controlled by Unified CVP, call admission control failures can be recovered with the proper Unified CCE Scripting configuration with the Unified CCE Router Requery feature enabled.

- H.323, SIP, or MGCP signaling traffic between the voice gateways and the centralized Unified CM servers will flow over the WAN. Proper QoS implementation on the WAN is critical, and signaling delays must be within tolerances listed in the latest version of the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

  http://www.cisco.com/go/ucsrnd
- Automated alternate routing (AAR) provides a mechanism to reroute calls through the PSTN or other network by using an alternate number when Unified CM blocks a call due to insufficient location bandwidth. For deployments with Unified CVP ingress call control, do not ena ble AAR, to allow Unified CVP Router Requery to take control of the call in the event of a failure or timeout. For deployments using IP-IVR, enable AAR to route the call through the PSTN or other network component using an alternative number.

Unified CCE: Unified CCE System PG

Because the deployment of contact center components is essentially the same as in other multi-site centralized call processing deployments, the same benefits and restrictions apply to Unified CCE deployed using the Unified CCE System PG.

Additionally, if Unified ICM/CCE pre-routing is used to interact with carriers and distribute calls to the voice gateways, translation routes for the NIC routing client to the Unified CCE System PG must be configured manually using the ConfigManager application on the Unified ICM/CCE Admin Workstation.

Unified CCE & IVR: Treatment and Queuing with Unified IP IVR

WAN bandwidth must be provisioned to support all calls that will be treated and queued at the central site.

IVR: Treatment and Queuing with Unified CVP

Unified CVP is not supported with a Unified CCE System PG. A separate VRU peripheral must be configured and deployed. This means that translation routes must be configured to transfer calls with call data between the peripherals. However, Unified CVP does provide benefits of queuing and treatment for callers at the remote distributed ingress voice gateways in this model because the calls do not have to cross the VoIP WAN for treatment in the centralized Unified IP IVR.

Using Unified CVP for treatment and queuing allows you to reduce the amount of voice bearer traffic traveling across the WAN. Unified CVP queues and treats calls on the remote gateways, thus eliminating the need to terminate the voice bearer traffic at the central site. WAN bandwidth must still be provisioned for transfers and conferences that involve agents at other locations.

Unified CCE: Unified CCE PG

Because the deployment of contact center components is essentially the same as in other multi-site centralized call processing deployments, the same benefits and restrictions apply to Unified CCE deployed using the Unified CCE PG.

Additionally, if Unified ICM/CCE pre-routing is used to interact with carriers and distribute calls to the voice gateways, translation routes must be configured for the NIC routing client using traditional Unified CCE with separate Unified CVP and Unified CM peripherals in the Unified ICM/CCE.

IVR: Treatment and Queuing with Unified IP IVR

WAN bandwidth must be provisioned to support all calls that will be treated and queued at the central site.

IVR: Treatment and Queuing with Unified CVP

Using Unified CVP for treatment and queuing allows you to reduce the amount of voice bearer traffic traveling across the WAN. Unified CVP queues and treats calls on the remote gateways, thus eliminating the need to terminate the voice bearer traffic at the central site. WAN bandwidth must still be provisioned for transfers and conferences that involve agents at other locations.

Unified CCE: Transfers

Intra-site or inter-site transfers using the VoIP WAN to send the RTP stream from one site to another occur basically the same way as a single-site transfer or a transfer in a deployment with centralized voice gateways.

An alternative to using the VoIP WAN for routing calls between sites is to use a carrier-based PSTN transfer service. These services allow the Unified CCE voice gateways to outpulse DTMF tones to instruct the PSTN to reroute (transfer) the call to another voice gateway location. Each site can be configured within the Unified ICM/CCE as a separate Agent Peripheral. The label then indicates whether a transfer is intra-site or inter-site, using Take Back and Transfer (*8) or Transfer Connect. These transfer tones are played in-band over the voice path and must be played from a recorded file in Unified IP IVR or outpulsed as digits from Unified CVP.

IPT: Multi-Site with Distributed Call Processing

Enterprises with multiple medium to large sites separated by large distances tend to prefer a distributed call processing model. In this model, each site has its own Unified CM cluster, treatment and queue points, PGs, and CTI Server. However, as with the centralized call processing model, sites could be

deployed with or without local voice gateways. Some deployments may also contain a combination of distributed voice gateways (possibly for locally dialed calls) and centralized voice gateways (possibly for toll-free calls) as well as centralized or distributed treatment and queue points.

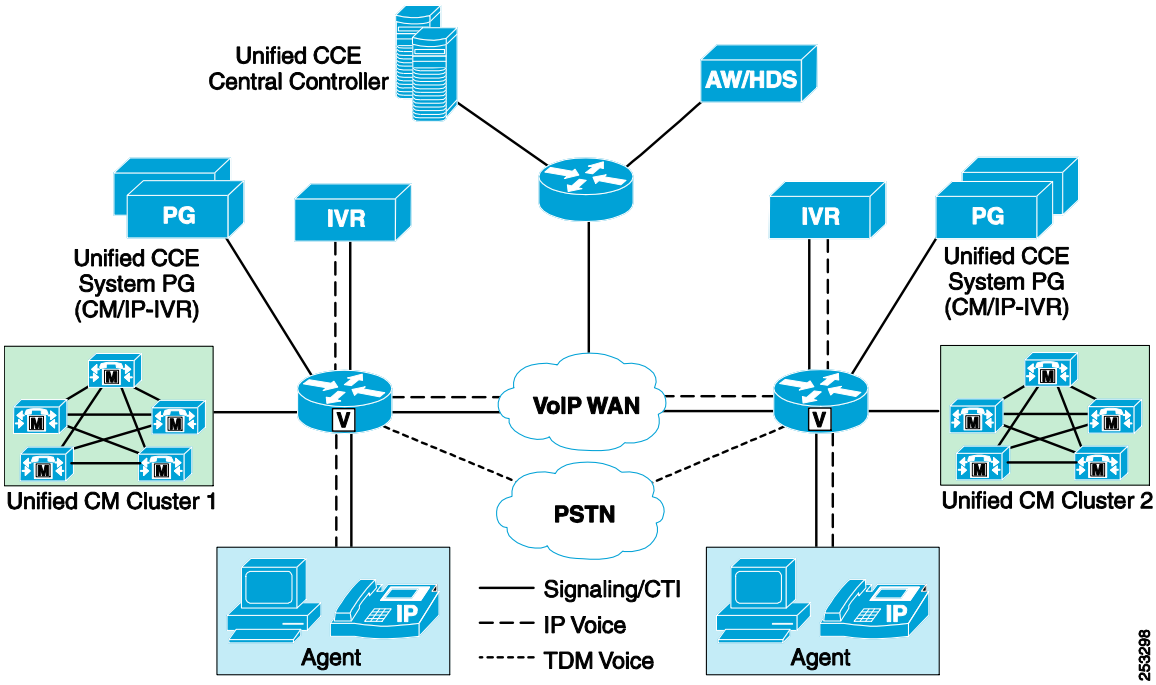
Regardless of how many sites are being deployed in this model, there will still be only one logical Unified CCE Central Controller. If the Unified CCE Central Controller is deployed with redundancy, sides A and B can be deployed side-by-side or geographically separated (remote redundancy). For details on remote redundancy, refer to the Unified ICM/CCE product documentation available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/tsd_products_support_series_home.html

Unified CCE: Distributed Voice Gateways with Treatment and Queuing Using Unified IP IVR

This deployment model is a good choice if the company has multiple medium to large sites. In this model, voice gateways with PSTN trunks terminate into each site. Just as in the centralized call processing model with distributed voice gateways, it might be desirable to limit the routing of calls to agents within the site where the call arrived (to reduce WAN bandwidth). An analysis of benefits from customer service levels versus WAN costs is required to determine whether limiting calls within a site is appropriate. Figure 2-11 illustrates this model using a traditional Unified CCE deployment with the Unified CCE System PG.

Figure 2-11 Multi-Site Deployment with Distributed Call Processing and Distributed Voice Gateways with Unified IP IVR



As with the previous models, many options are possible. The number and type of Unified CCE Servers, Unified CM servers, and Unified IP IVR servers can vary. LAN/WAN infrastructure, voice gateways, PSTN trunks, redundancy, and so forth are also variable within this deployment model. Central processing and gateways may be added for self-service, toll-free calls and support for smaller sites. In addition, the use of a pre-routing PSTN Network Interface Controller (NIC) is also an option.

Advantages

- Scalability — Each independent site can scale up to the maximum number of supported agents per Unified CM cluster, and there is no software limit to the number of sites that can be combined by the Unified CCE Central Controller to produce a single enterprise-wide contact center, provided that the total concurrent agent count is less than the maximum supported agent count in a Unified CCE System. For scalability and sizing information, see Chapter 10, "Sizing Unified CCE Components and Servers".
- All or most VoIP traffic can be contained within the LAN of each site, if desired. The QoS WAN shown in Figure 2-11 would be required for voice calls to be transferred across sites. Use of a PSTN transfer service (for example, Take Back and Transfer or Transfer Connect) could eliminate that need. If desired, a small portion of calls arriving at a particular site can be queued for agent resources at other sites to improve customer service levels.
- Unified ICM/CCE pre-routing can be used to load-balance calls based on agent or Unified IP IVR port availability to the best site to reduce WAN usage for VoIP traffic.
- Failure at any one site has no impact on operations at another site.
- Each site can be sized according to the requirements for that site
- The Unified CCE Central Controller provides centralized management for configuration of routing for all calls within the enterprise.
- The Unified CCE Central Controller provides the capability to create a single enterprise-wide queue.
- The Unified CCE Central Controller provides consolidated reporting for all sites.

Best Practices

- The PG, Unified CM cluster, and Unified IP IVR must be co-located at the contact center site.
- The communication link from the Unified CCE Central Controller to the PG must be sized properly and provisioned for bandwidth and QoS. (For details, refer to Chapter 12, "Bandwidth Provisioning and QoS Considerations".)
- Gatekeeper-based or RSVP Agent-based call admission control could be used to reroute calls between sites over the PSTN when WAN bandwidth is not available. It is best to ensure that adequate WAN bandwidth exists between sites for the maximum amount of calling that can occur.
- If the communication link between the PG and the Unified CCE Central Controller is lost, then all contact center routing for calls at that site is also lost. Therefore, it is important to implement a fault-tolerant WAN. Even when a fault-tolerant WAN is implemented, it is important to identify contingency plans for call treatment and routing when communication is lost between the Unified CCE Central Controller and PG. For example, in the event of a lost Unified CCE Central Controller connection, the Unified CM CTI route points could send the calls to Unified IP IVR ports to provide basic announcement treatment or to invoke a PSTN transfer to another site. Another alternative is for the Unified CM cluster to route the call to another Unified CM cluster that has a PG with an active connection to the Unified CCE Central Controller. For more information on these options, refer to Chapter 3, "Design Considerations for High Availability".
- While two intercluster call legs for the same call will not cause unnecessary RTP streams, two separate call signaling control paths will remain intact between the two clusters (producing logical hairpinning and reducing the number of intercluster trunks by two). Consider the percentage of inter-site transfers when sizing intercluster trunks capacities.
- Latency between Unified CCE Central Controllers and remote PGs cannot exceed 200 ms one way (400 ms round-trip).

Treatment and Queuing

Initial call queuing is done on a Unified IP IVR co-located with the voice gateways, so no transcoding is required. When a call is transferred and subsequent queuing is required, perform the queuing on a Unified IP IVR at the site where the call is currently being processed. For example, if a call comes into Site 1 and gets routed to an agent at Site 2, but that agent needs to transfer the call to another agent whose location is unknown, queue the call to a Unified IP IVR at Site 2 to avoid generating another intercluster call. A second intercluster call would be made only if an agent at Site 1 was selected for the transfer. The RTP flow at this point would be directly from the voice gateway at Site 1 to the agent's phone at Site 1. However, the two Unified CM clusters would still logically see two calls in progress between the two clusters.

Transfers

Transfers within a site function just like a single-site transfer. Transfers between Unified CM clusters use either the VoIP WAN or a PSTN service.

If the VoIP WAN is used, sufficient intercluster trunks must be configured. An alternative to using the VoIP WAN for routing calls between sites is to use a PSTN transfer service. These services allow the Unified CCE voice gateways to outpulse DTMF tones to instruct the PSTN to reroute (transfer) the call to another voice gateway location. Another alternative is to have the Unified CM cluster at Site 1 make an outbound call back to the PSTN. The PSTN would then route the call to Site 2, but the call would use two voice gateway ports at Site 1 for the remainder of the call.

Unified CCE: Unified CCE System PG

The Unified CCE System PG acts as a single peripheral that joins the Unified CM and Unified IP IVR peripherals of former versions to simplify installation, configuration, and routing. In this model, the PGs at the remote sites can be installed as Unified CCE System PGs to combine the Unified IP IVR and Unified CM peripherals under a single logical PG instance and peripheral.

This model is perhaps more typical of outsourcers that would set up a call center specifically for a single client and deploy it as a Unified CCE System PG to allow their client company to connect their Unified CCE Enterprise system to the outsourcer Unified CCE System PG with the Unified CCE Gateway PG, as they would any outsourced ACD.

Unified CCE: Unified CCE PG

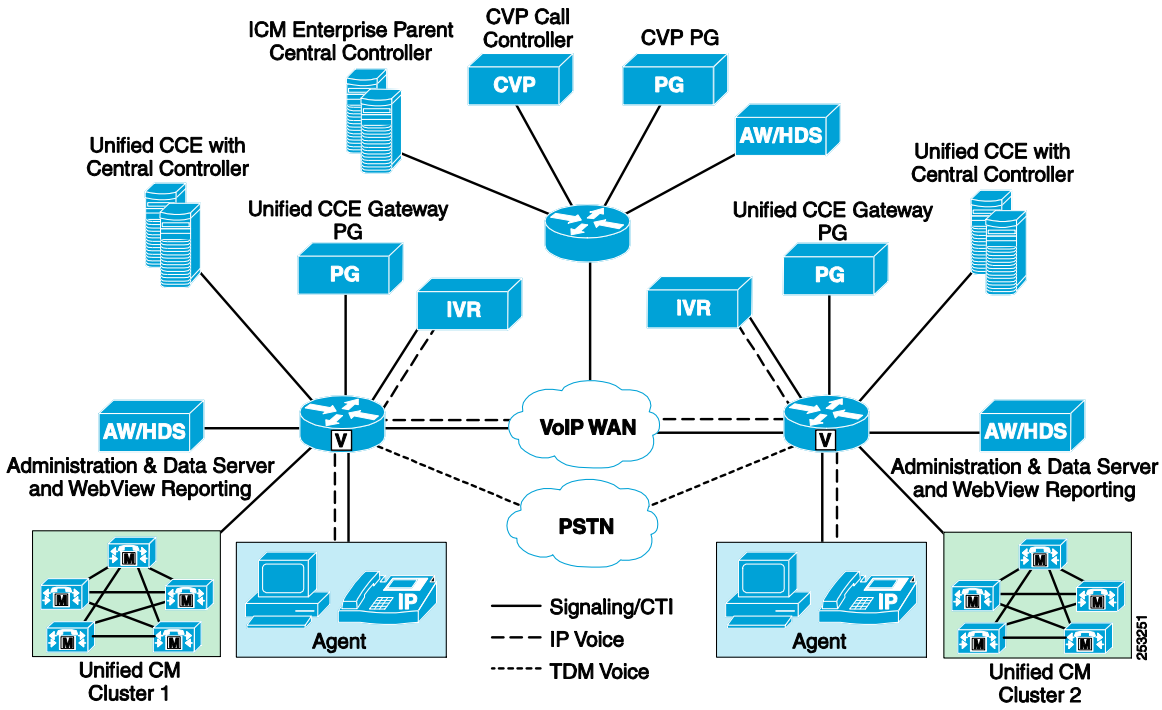
This model, as designed with multiple remote locations, is more suited for the traditional Unified CCE design with multiple distributed peripheral gateways. The system could be deployed with the Generic PG or both Unified CM and Unified IP IVR PGs at the sites; however, the new Unified CCE System PG that combines both of these peripherals into a single peripheral for routing and reporting under the traditional model might be easier for new deployments of this solution. Existing customers upgrading to Unified CCE 8.0 may stay on their existing Generic PG or multi-PG model.

Alternative: Parent/Child

The Unified ICM Enterprise (parent) and Unified CCE (child) model is appropriate alternative deployment to provide local, distributed call processing with a local Unified CM and Unified CCE at each site (child), controlled under a centralized Unified ICM Enterprise parent for enterprise-wide

routing, reporting, and call control. This model has the advantage of being more tolerant of WAN outages, and each site is completely survivable. Figure 2-12 shows this same model deployed using the parent/child model.

Figure 2-12 Multi-Site Deployment with Distributed Call Processing and Parent/Child



In this design, there is a parent Unified ICM Enterprise system deployed with Unified CVP and its own Administration & Data server. At each distributed site, there is a complete Unified CCE deployment consisting of Central Controller on one or more servers. There is also a local Administration & Data Server for Unified CCE to perform configuration, scripting, and reporting tasks for that specific site. There is a Unified CCE Gateway PG that connects Unified CCE to the Unified ICM parent, and it is part of the Peripheral Gateways deployed on the parent Unified ICM. An optional deployment for the Unified CCE Gateway PG is to co-locate it with the Unified CCE System PG, while adhering to the following guidelines:

If the Unified CCE Gateway PG and Unified CCE System PG Instance Numbers are the same, then the PG number for the Unified CCE Gateway PG and Unified CCE System PG must be different.

If the Unified CCE Gateway PG and Unified CCE System PG Instance Numbers are different, then the PG number for the Unified CCE Gateway PG and the Unified CCE System PG may be the same.

No additional PGs (such as a VRU PG or MR PG) can be added to this Server.

The server co-resident Unified CCE Gateway PG and Unified CCE System PG is not supported with a Unified System CCE deployment.

For scalability limits of the co-resident Unified CCE Gateway PG and Unified CCE System PG, refer to "Sizing Unified CCE Components and Servers" for additional details.

In this design, the local Unified CCE deployments act as their own local IP ACDs, with no visibility to any of the other sites in the system. Users at Site 1 cannot see any of the calls or reports from Site 2 in this model. Only the Unified ICM Enterprise parent system has visibility to all activity at all sites connected to the Unified ICM Enterprise system.

The Unified CVP at the Unified ICM parent site is used to control the calls coming into the distributed sites, providing local call queuing and treatment in the VoiceXML Browser in the voice gateway. When configuring the Unified ICM parent CVP to use Unified CVP Router Requery, to take control of the call in the event of a failure or answer timeout, the child Unified CCE cannot terminate the ingress call to a child Unified CVP or a child Unified IP IVR. The local Unified IP IVR servers are used only for a local backup if the connection from these voice gateways is lost to the parent Unified CVP Call Control server. The local Unified IP IVR also provides local queue treatment for calls that are not answered by the local agents (RONA), rather than sending the call back to the Unified CVP to be re-queued.

The child Unified CCE deployments can also transfer calls across the system between the sites using Unified ICM post-routing by the Unified CCE Gateway PG. The Unified CCE Gateway PG allows the child Unified CCE to ask the Unified ICM to transfer a call to the best agent at another site or to queue it centrally for the next available agent.

Unlike traditional Unified CCE models with distributed Unified CM Peripheral Gateways, the parent/child model provides for complete local redundancy at the contact center site. The local Unified CCE will take over call processing for inbound calls from the Unified CVP gateways and provide local call queuing and treatment in the local Unified IP IVR. This is an excellent design for call center sites that require complete redundancy or 100% up-time and that cannot be down because of a WAN failure.

This design is a good approach for customers who have Unified ICM already installed with their TDM ACD platforms and who want either to add new sites with Unified CCE or to convert an existing site to Unified CCE. It allows the Unified ICM to continue performing enterprise-wide routing and reporting across all of the sites while inserting new Unified CCE technology on a site-by-site basis.

Note also that Unified CVP could be at both the parent and child. This is virtually identical to Unified CVP at the parent and IP-IVR at the child from a call flow perspective. One key difference, though, is that information on queued calls at the child Unified CVP will not be available at the parent (through the Unified CCE Gateway PG), as is the case if IP-IVR is used. This means that minimum expected delay (MED) over services cannot be used.

Advantages

- Unified CVP provides a virtual network queue across all the distributed sites controlled by the parent Unified ICM. The parent Unified ICM has visibility into all the distributed sites and will send the call to the next available agent from the virtual queue.
- Each distributed site can scale up to the maximum number of supported agents on a single Unified CCE deployment. Multiple Unified CCE Central Controllers can be connected to a single Unified CM cluster to scale up to the maximum number of supported agents per cluster. The Unified CCE systems are connected to the parent Unified ICM using the Unified CCE Gateway PG on the parent Unified ICM, which can scale up to the maximum number of supported agents per parent Unified ICM Enterprise system.
- All or most VoIP traffic can be contained within the LAN of each site, if desired. The QoS WAN shown in Figure 2-12 would be required for voice calls to be transferred across sites. Use of a PSTN transfer service (for example, Take Back and Transfer or Transfer Connect) could eliminate that need. If desired, a small portion of calls arriving at a particular site can be queued for agent resources at other sites to improve customer service levels.
- Unified ICM pre-routing can be used to load-balance calls based on agent or Unified CVP session availability and to route calls to the best site to reduce WAN usage for VoIP traffic.
- Failure at any one site has no impact on operations at another site.
- Each site can be sized according to the requirements for that site
- The parent Unified ICM Central Controller provides centralized management for configuration of routing for all calls within the enterprise.

- The parent Unified ICM Central Controller provides the capability to create a single enterprise-wide queue.
- The parent Unified ICM Central Controller provides consolidated reporting for all sites.

Disadvantages

- Server count — The number of servers that are required to manage the parent/child model is usually higher due to the increased number of software components (additional Unified CCE Gateway PGs required if co-locating with Unified CCE System PG is not an option, additional Central Controller for each child, and so forth).

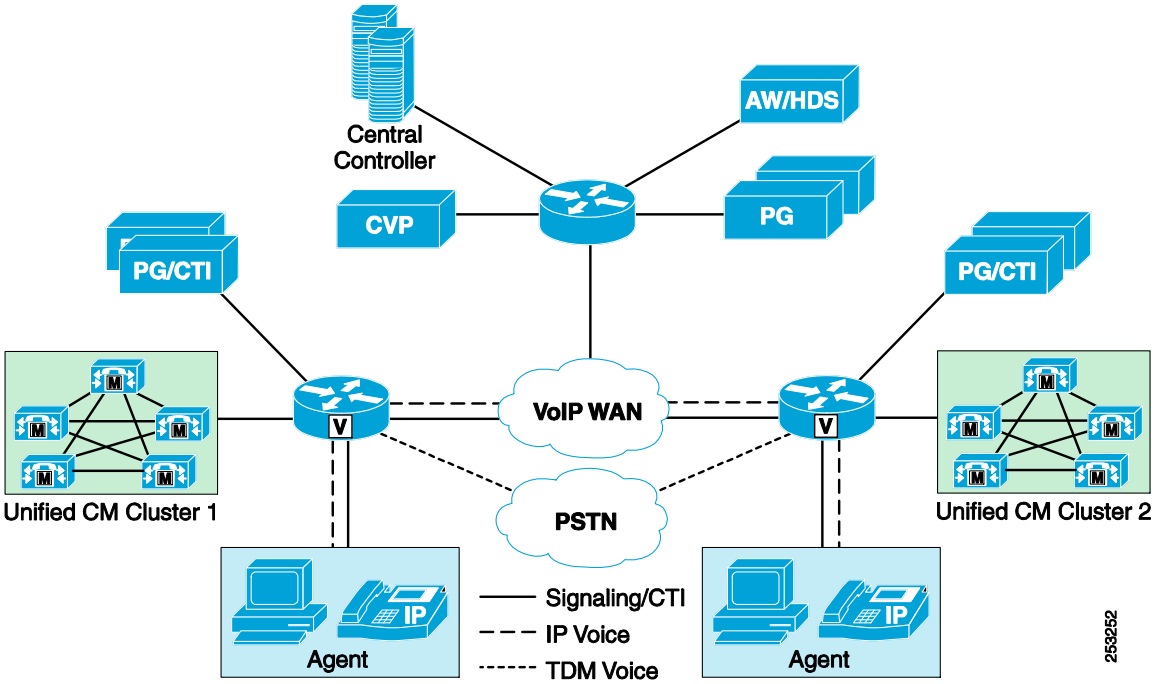
Best Practices

- Co-locate the Unified CCE Gateway PG, Unified CM cluster, Unified IP IVR, and Unified CCE (if possible) be co-located at the contact center site.
- The communication link from the parent Unified ICM Central Controller to the Unified CCE Gateway PG must be sized properly and provisioned for bandwidth and QoS. (For details, refer to the Chapter 12, "Bandwidth Provisioning and QoS Considerations".)
- Gatekeeper-based or RSVP agent-based call admission control could be used to reroute calls between sites over the PSTN when WAN bandwidth is not available. It is best to ensure that adequate WAN bandwidth exists between sites for the maximum amount of calling that can occur.
- If the communication link between the Unified CCE Gateway PG and the parent Unified ICM Central Controller is lost, then all contact center routing for calls at that site is put under control of the local Unified CCE. Unified CVP-controlled ingress voice gateways would have survivability TCL scripts to redirect inbound calls to local Unified CM CTI route points, and the local Unified IP IVR would be used to handle local queuing and treatment during the WAN outage. This is a major feature of the parent/child model to provide complete local survivability for the call center. For more information, see Chapter 3, "Design Considerations for High Availability".
- While two intercluster call legs for the same call will not cause unnecessary RTP streams, two separate call signaling control paths will remain intact between the two clusters (producing logical hairpinning and reducing the number of intercluster trunks by two). Consider the percentage of inter-site transfers when sizing intercluster trunks capacities.
- Latency between parent Unified ICM Central Controllers and remote Unified CCE Gateway PGs must not exceed 200 ms one way (400 ms round-trip).

IVR: Distributed Voice Gateways with Treatment and Queuing Using Unified CVP

This deployment model is the same as the previous model, except that Unified CVP is used instead of Unified IP IVR for call treatment and queuing. In this model, voice gateways with PSTN trunks terminate into each site. Just as in the centralized call processing model with distributed voice gateways, it might be desirable to limit the routing of calls to agents within the site where the call arrived (to reduce WAN bandwidth). Call treatment and queuing can also be achieved at the site where the call arrived, further reducing the WAN bandwidth needs. Figure 2-13 illustrates this model using a traditional Unified CCE deployment.

Figure 2-13 Multi-Site Deployment with Distributed Call Processing and Distributed Voice Gateways with Unified CVP



As with the previous models, many options are possible. The number and type of Unified CCE Servers, Unified CM servers, and Unified CVP servers can vary. LAN/WAN infrastructure, voice gateways, PSTN trunks, redundancy, and so forth, are also variable within this deployment model. Central processing and gateways may be added for self-service, toll-free calls and support for smaller sites. In addition, the use of a pre-routing PSTN Network Interface Controller (NIC) is also an option.

Advantages

- Unified CVP Servers can be located either centrally or remotely. Call treatment and queuing will still be distributed, executing on the local gateway, regardless of Unified CVP server location. Unified CVP is shown centrally located in Figure 2-13.
- For information on the number of agents supported per PG and across the entire system, see Chapter 10, "Sizing Unified CCE Components and Servers".
- All or most VoIP traffic can be contained within the LAN of each site, if desired. The QoS WAN would be required for voice calls to be transferred across sites. Use of a PSTN transfer service (for example, Takeback N Transfer) could eliminate that need. If desired, a small portion of calls arriving at a particular site can be queued for agent resources at other sites to improve customer service levels.
- Unified CCE pre-routing can be used to load-balance calls and route them to the best site to reduce WAN usage for VoIP traffic.
- Failure at any one site has no impact on operations at another site.
- Each site can be sized according to the requirements for that site.
- The Unified CCE Central Controller provides centralized management for configuration of routing for all calls within the enterprise.
- The Unified CCE Central Controller provides the capability to create a single enterprise-wide queue.

- The Unified CCE Central Controller provides consolidated reporting for all sites.

Best Practices

- The Unified CM PG and Unified CM cluster must be co-located. The Unified CVP PG and Unified CVP servers must be co-located.
- The communication link from the Unified CCE Central Controller to PG must be properly sized and provisioned for bandwidth and QoS. Cisco provides a partner tool called the *VRU Peripheral Gateway to Unified ICM Central Controller Bandwidth Calculator* to assist in calculating the VRU PG-to-Unified ICM/CCE bandwidth requirement. This tool is available (with valid Cisco Partner login authentication) at

  http://www.cisco.com/web/partners/sell/technology/ipc/integrated-solutions/customer_contact_center.html
- If the communication link between the PG and the Unified CCE Central Controller is lost, then all contact center routing for calls at that site is lost. Therefore, it is important to implement a fault-tolerant WAN. Even when a fault-tolerant WAN is implemented, it is important to identify contingency plans for call treatment and routing when communication is lost between the Unified CCE Central Controller and PG.
- Latency between Unified CCE Central Controllers and remote PGs must not exceed 200 ms one way (400 ms round-trip)

IVR: Treatment and Queuing

Unified CVP queues and treats calls on the remote gateways, eliminating the need to terminate the voice bearer traffic at the central site. Unified CVP servers may be located at the central site or distributed to remote sites. WAN bandwidth must still be provisioned for transfers and conferences that involve agents at other locations.

Unlike Unified IP IVR, with Unified CVP the call legs are torn down and reconnected, avoiding signaling hairpins. With Unified IP IVR, two separate call signaling control paths will remain intact between the two clusters (producing logical hairpinning and reducing the number of intercluster trunks by two).

Transfers

Transfers within a site function just like a single-site transfer. Transfers between Unified CM clusters use either the VoIP WAN or a PSTN service.

If the VoIP WAN is used, sufficient intercluster trunks must be configured. An alternative to using the VoIP WAN for routing calls between sites is to use a PSTN transfer service. These services allow the Unified CCE voice gateways to outpulse DTMF tones to instruct the PSTN to reroute (transfer) the call to another voice gateway location. Another alternative is to have the Unified CM cluster at Site 1 make an outbound call back to the PSTN. The PSTN would then route the call to Site 2, but the call would use two voice gateway ports at Site 1 for the remainder of the call.

Unified CCE: Unified CCE System PG

The Unified CCE System PG is not a good fit for this model because it does not support Unified CVP for queuing, and the IVR PIMs on the Unified CCE System PG would go unused.

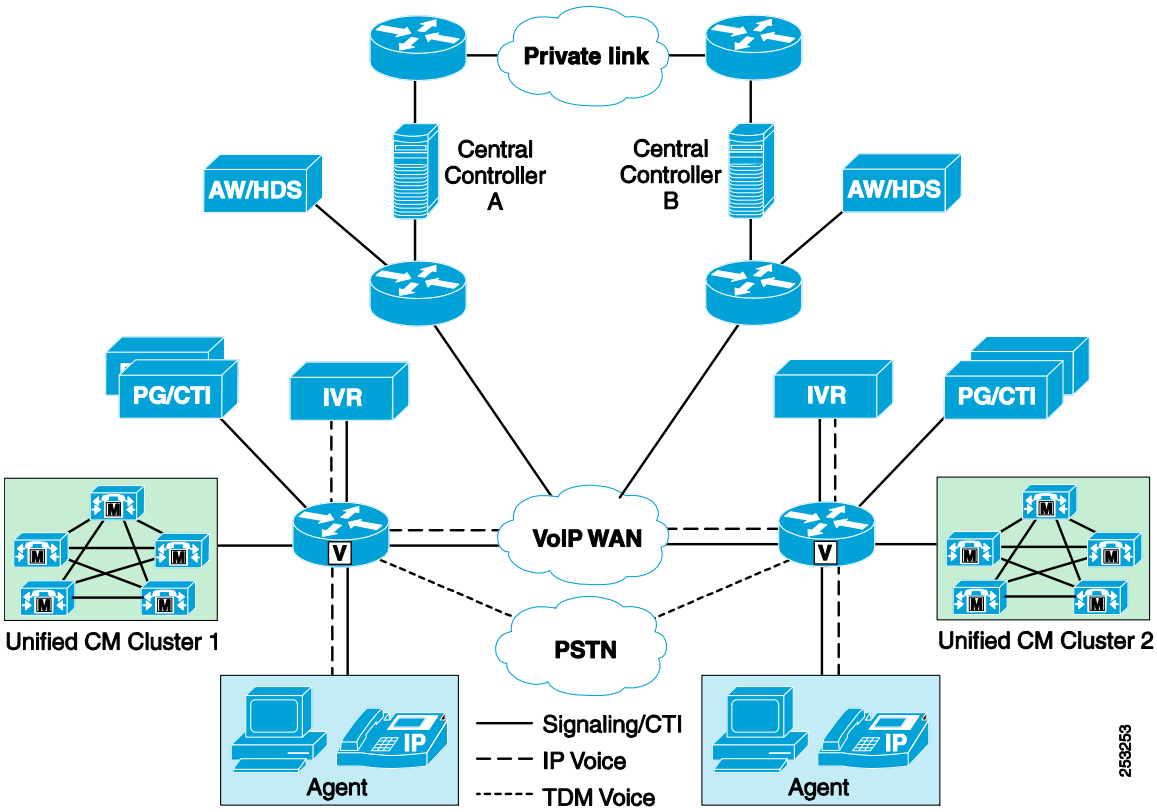
Unified CCE: Unified CCE PG

The Unified CCE PG is the required PG for this deployment model.

Unified CCE: Distributed Unified CCE Option with Distributed Call Processing Model

Figure 2-14 illustrates this deployment model.

Figure 2-14 Distributed Unified CCE Option Shown with Unified IP IVR



Advantages

The primary advantage of the distributed Unified CCE option is the redundancy gained from splitting the Unified CCE Central Controller between two redundant sites.

Best Practices

- Unified CCE Central Controllers (Routers and Loggers) require a separate network path or link to carry the private communications between the two redundant sites. In a non-distributed Unified CCE model, the private traffic usually traverses an Ethernet crossover cable or LAN connected directly between the side A and side B Unified CCE Central Controller components. In the distributed Unified CCE model, the private communications between the A and B Unified CCE components travel across a dedicated link with at least as much bandwidth as a T1 line.
- Latency across the private separate link must not exceed 100 ms one way (200 ms round-trip), but 50 ms (100 ms round-trip) is preferred.

- Latency between Unified CCE Central Controllers and remote PGs must not exceed 200 ms one way (400 ms round-trip).
- The private link cannot traverse the same path as public traffic. The private link must have path diversity and must reside on a link that is completely path-independent from Unified CCE public traffic. This link is used as part of the system fault tolerant design. For more information, see Chapter 3, "Design Considerations for High Availability".
- The redundant centralized model is explored in the next section on IPT: Clustering Over the WAN , page 2-32.
- Automated alternate routing (AAR) provides a mechanism to reroute calls through the PSTN or other network by using an alternate number when Unified CM blocks a call due to insufficient location bandwidth. For deployments with Unified CVP ingress call control, do not enable AAR, to allow Unified CVP Router Requery to take control of the call in the event of a failure or timeout. For deployments using IP-IVR, enable AAR to route the call through the PSTN or other network component using an alternative number.

IPT: Clustering Over the WAN

As part of the centralization of call processing, many customers prefer to combine the redundancy of the distributed Unified CM call processing model with the simplicity of having a single Unified CM cluster for a single dial plan and voice system to administer. This combination of models provides for a single Unified CM cluster with its subscriber servers split across data center locations to provide a single cluster with multiple distributed call processing servers for a highly available and redundant design, known as clustering over the WAN.

Unified CM clustering over the WAN may also be used with Unified CCE for contact centers to allow full agent redundancy in the case of a data center (central site) outage. Implementation of clustering over the WAN for Unified CCE does have several strict requirements that differ from other models. Bandwidth between central sites for Unified CCE public and private traffic, Unified CM intra-cluster communication signaling (ICCS), and all other voice-related media and signaling must be properly provisioned with QoS enabled. The WAN between central sites must be highly available (HA) with redundant links and redundant routers.

Advantages

- No single point of failure, including loss of an entire central site.
- Cisco Unified Mobile Agents (Remote Agent) require no reconfiguration to remain fully operational in case of site or link outage. When outages occur, agents and agent devices dynamically switch to the redundant site.
- Central administration for both Unified CCE and Unified CM.
- Reduction of servers for distributed deployment.

Best Practices

- Deploy a minimum of three WAN links for systems that employ clustered over the WAN. Deploy at least two links for the highly available network that carries the Unified CCE public traffic (see Figure 2-15). Use aseparate WAN link for the Unified CCE private traffic (see Figure 2-16). If QoS and bandwidth are configured correctly (see the guidelines in Chapter 12, "Bandwidth Provisioning and QoS Considerations"), these WAN links can be converged with other corporate traffic as long as the private and public traffic are not carried over the same link. Carry the Unified CM ICCS traffic over the highly available network used by the Unified CCE public communications.

Figure 2-15 Highly Available WAN Network for the Unified CCE Public Traffic

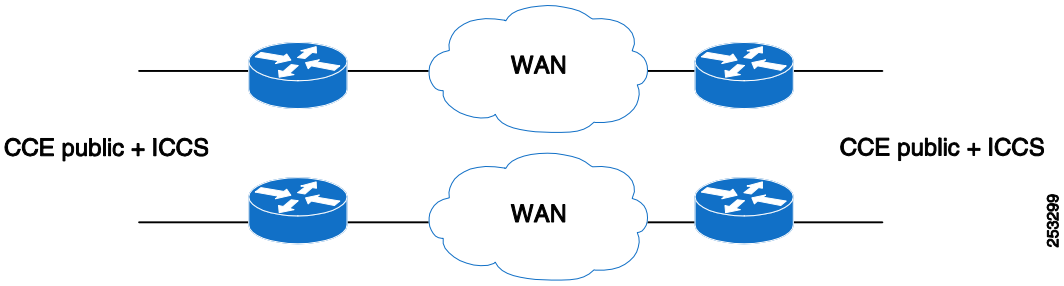
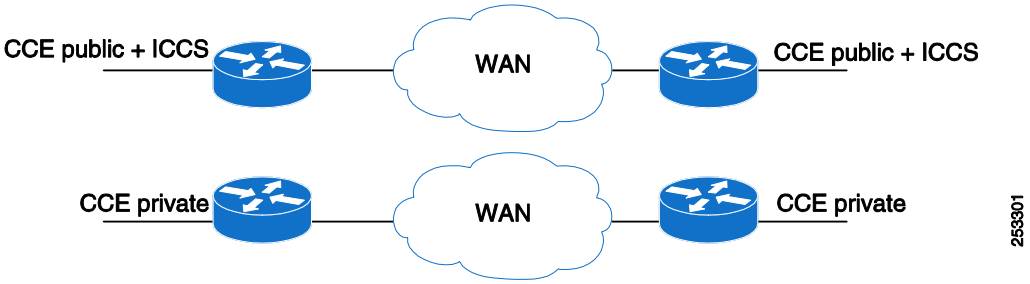


Figure 2-16 Separate WAN link for Unified CCE Private Traffic

- It is possible to deploy Unified CCE clustering over the WAN with two links if the following rules are applied:
  - During normal operations, the Unified CCE public and private traffic must be carried over separate links; they must not be carried over the same link.
  - Carry the Unified CM traffic over the Unified CCE public link in normal conditions (see Figure 2-17).
  - Two routers are required on each side of the WAN for redundancy; connected these to different WAN links.

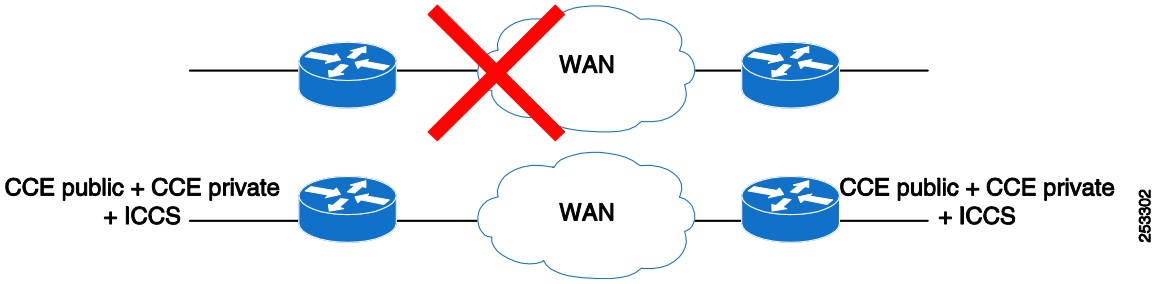
Figure 2-17 Network Architecture Under Normal Operations



- In case of network failure, configure the WAN link that carries the Unified CCE public traffic to fail-over to the other link that carries the Unified CCE private traffic (see Figure 2-18). Temporarily allow the Unified CM ICCS traffic to fail-over to the private link. This prevents situations where a CTI Manager that connects to the active Agent PG loses its WAN connection to the Unified CM node to which the agent phones are registered. Restore the link as fast as possible so that the public and private Unified CCE traffic are carried over separate links. If the redundant link that carries the Unified CCE private traffic also fails, Unified CCE instability and data loss can occur, including the corruption of one Logger database, and manual intervention could be required. That is why it is very important to actively monitor any network failure at all times.

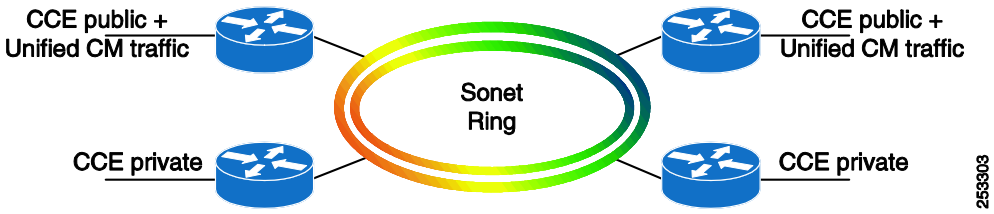
  The links must also be sized correctly in order to accommodate this failover situation where the private link carries the entire WAN traffic, including the public and ICCS traffic. QoS and bandwidth must be configured according to the guidelines in Chapter 12, "Bandwidth Provisioning and QoS Considerations".

Figure 2-18 Network Architecture After Failure of the Unified CCE Public Network



- It is also possible to allow the private link to fail-over to the public link. However, if the total failover latency takes more than 500 ms (five times the TCP keepalive interval of 100 ms), the Unified CCE system considers the private link to be down. If the public link is also down, Unified CCE instability and data loss can occur, including the corruption of one Logger database, and manual intervention could be required. The total failover latency typically includes the round-trip transmission latency, the routing protocol convergence delay, the HSRP convergence delay if applicable, queuing and packetization delays, and any other delay that would be applicable. If the total failover latency is higher than 500 ms, or if you suspect possible recurrent network flapping,deploy three WAN links and keeping the private traffic separate from the public traffic at all times. Also, the links must be sized correctly in order to accommodate this failover situation where the public link carries the entire WAN traffic, including the private and ICCS traffic. Restore the link must be restored as fast as possible so that the public and private Unified CCE traffic are carried over separate links.
- If QoS and bandwidth are configured correctly (see the guidelines in Chapter 12, "Bandwidth Provisioning and QoS Considerations" for more details), these WAN links can be converged with other corporate traffic.
- With a SONET fiber ring, which is highly resilient and has built-in redundancy, the public and private traffic can be carried over the same SONET ring under normal operations or following a network failover. A separate link for the private traffic is not required in this case. Also, two routers are required on each side of the WAN for redundancy. Under normal operations, use one router for the Unified CCE public traffic and use the other router for the Unified CCE private traffic. (See Figure 2-19.) The other rules described in this section also apply.

Figure 2-19 Network Architecture Based on a SONET Ring Under Normal Operations



- The highly available (HA) WAN between the central sites must be fully redundant with no single point of failure. (For information regarding site-to-site redundancy options, refer to the WAN infrastructure and QoS design guides available at http://www.cisco.com/go/designzone.) In case of partial failure of the highly available WAN, the redundant link must be capable of handling the full central-site load with all QoS parameters. For more information, see the section on Bandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20.
- A highly available (HA) WAN using point-to-point technology is best implemented across two separate carriers, but this is not necessary when using a ring technology.

- Latency requirements across the highly available (HA) WAN must meet the current Cisco Unified Communications requirements for clustering over the WAN. Currently, a maximum latency of 40 ms one way (80 ms round-trip) is allowed with Unified CM 6.1 or later releases. With prior versions of Unified CM, the maximum latency is 20 ms one way. For full specifications, refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

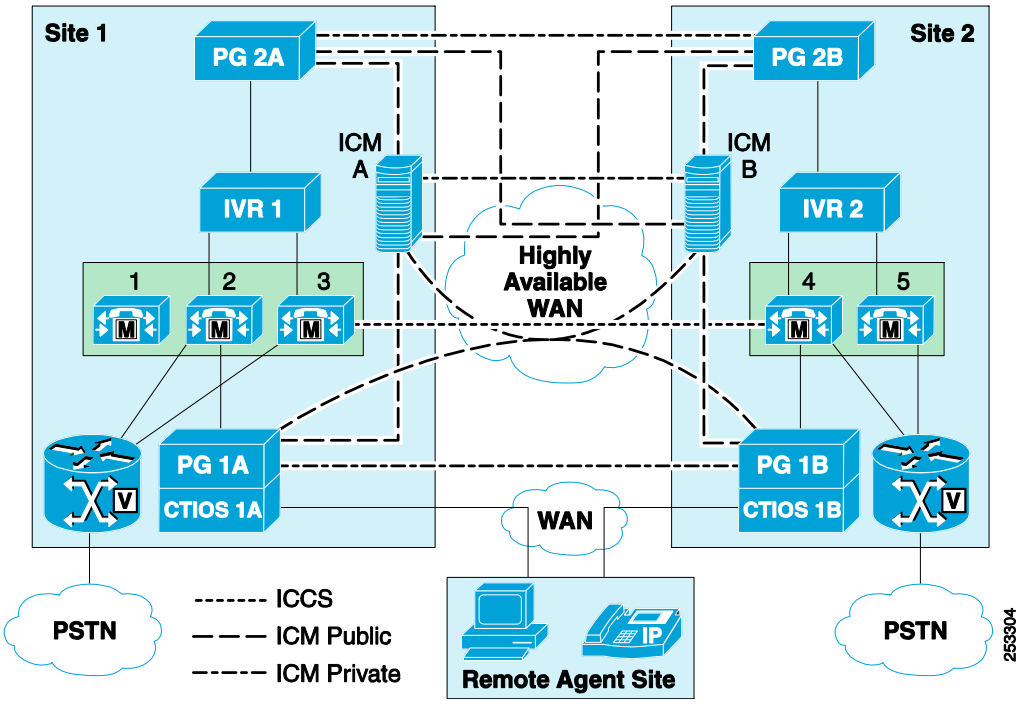
  http://www.cisco.com/go/ucsrnd

- Unified CCE latency requirements can be met by conforming to Cisco Unified Communications requirements. However, the bandwidth requirements for Unified CM intra-cluster communications differ between Unified CCE and Cisco Unified Communications. For more information, see the section on Bandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20.
- Bandwidth requirements across the highly available (HA) WAN include bandwidth and QoS provisioning for (see Bandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20):
  - Unified CM intra-cluster communication signaling (ICCS)
  - Communications between Unified CCE Central Controllers
  - Communications between Unified CCE Central Controller and PG
  - Communications between CTI Object Server (CTI OS) and CTI Server, if using CTI OS
- Deploy separate dedicated link(s) for Unified CCE private communications between Unified CCE Central Controllers Side A and Side B and between PGs Side A and Side B to ensure path diversity. Path diversity is required due to the architecture of Unified CCE. Without path diversity, the possibility of a dual (public communication and private communication) failure exists. If a dual failure occurs even for a moment, Unified CCE instability and data loss can occur, including the corruption of one Logger database. The separate link(s) for Unified CCE private communications can be converged with other corporate traffic if QoS and bandwidth are configured correctly, but they cannot be converged with the Unified CCE public traffic.
- The separate private link(s) may be either two links (one for Central Controller private traffic and one for Unified CM PG private traffic) or one converged link containing both Central Controller and PG private traffic. SeeSite-to-Site Unified CCE Private Communications Options , for more information.
- Separate paths must exist from agent sites to each central site. Both paths must be capable of handling the full load of signaling, media and other traffic if one path fails. These paths may reside on the same physical link from the agent site, with a WAN technology such as Frame Relay using multiple permanent virtual circuits (PVCs).
- The minimum cluster size using Unified IP IVR as the treatment and queuing platform is 5 nodes (publisher plus 4 subscribers). This minimum is required to allow Unified IP IVR at each site to have redundant connections locally to the cluster without traversing the WAN. JTAPI connectivity between Unified CM and Unified IP IVR is not supported across the WAN in this model. Local gateways also will need local redundant connections to Unified CM.
- The minimum cluster size using Unified CVP as the treatment and queuing platform is 3 nodes (publisher plus 2 subscribers). However, a deployment with 5 nodes is preferable, especially if there are phones (either contact center or non-contact center) local to the central sites, central gateways, or central media resources, which would require local failover capabilities.
- In a deployment with clustering over the WAN, the VRU PG could connect to a local IP IVR or Unified CVP, or to a redundant IP IVR or Unified CVP across the WAN. For information on bandwidth requirements, refer to Chapter 12, “Bandwidth Provisioning and QoS Considerations”.
- In a deployment with clustering over the WAN, the Unified CM PG must be on the same LAN segment with the CTI Manager to which it is connected.

Centralized Voice Gateways with Centralized Call Treatment and Queuing Using Unified IP IVR

In this model, the voice gateways are located in the central sites. Unified IP IVR is centrally located and used for treatment and queuing on each side. Figure 2-20 illustrates this model.

Figure 2-20 Centralized Voice Gateways with Centralized Call Treatment and Queuing Using Unified IP IVR



Advantages

- Component location and administration are centralized.
- Calls are treated and queued locally, eliminating the need for queuing across a WAN connection.

Best Practices

- WAN connections to agent sites must be provisioned with bandwidth for voice as well as control and CTI. See Bandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20, for more information.
- A local voice gateway might be needed at remote sites for local out-calling and 911. For more information, refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

  http://www.cisco.com/go/ucsrnd
- Central site outages would include loss of half of the ingress gateways, assuming a balanced deployment. Gateways and IVRs must be scaled to handle the full load in both sites if one site fails.
- Carrier call routing must be able to route calls to the alternate site in the case of a site or gateway loss. Pre-routing may be used to balance the load, but it will not be able to prevent calls from being routed to a failed central site. Pre-routing is best deployed only as a last resort.

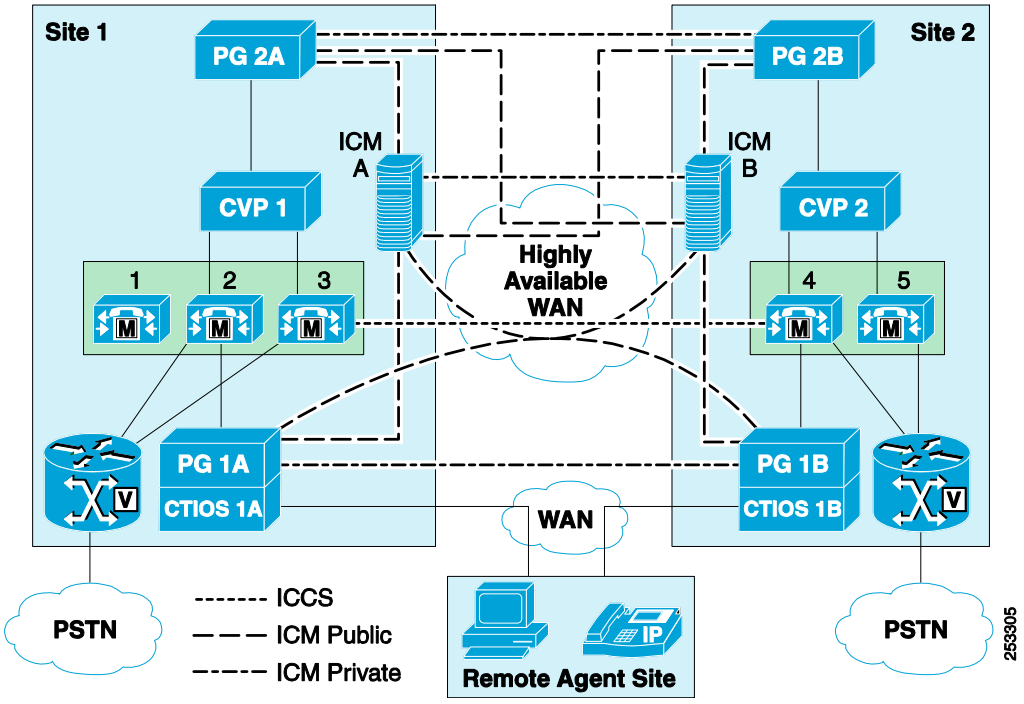
Clustering Over the WAN with Unified CCE System PG

Clustering over the WAN with Unified CCE System PG is now supported. However, due to the fact that a single Unified CCE System peripheral is controlling all of the Unified IP IVRs and the Unified CM, the load-balancing of calls between Unified IP IVRs does not take into account which site the call came into; it simply distributes the calls to whichever Unified IP IVR is least loaded. This means that calls coming into Site A might be treated by a Unified IP IVR in Site B. Additionally, both the A-side and B-side Unified CCE System PG know about all of the Unified IP IVRs. PIM activation logic will determine if the A-side or the B-side PIM will connect to each of the Unified IP IVRs. This means that the PG at site A might connect to the Unified IP IVR at site B. This means traffic might not be sent optimally over the WAN. In this model, make sure the WAN is sized for proper operation given this fact. To avoid this bandwidth overhead, you can optionally consider Clustering Over the WAN deployments with Unified CVP in place of Unified IP-IVR.

Centralized Voice Gateways with Centralized Call Treatment and Queuing Using Unified CVP

In this model, the voice gateways are VoiceXML gateways located in the central sites. Unified CVP is centrally located and used for treatment and queuing. Figure 2-21 illustrates this model.

Figure 2-21 Centralized Voice Gateways with Centralized Call Treatment and Queuing Using Unified CVP



Advantages

- Component location and administration are centralized.
- Calls are treated and queued locally, eliminating the need for queuing across a WAN connection.
- There is less load on Unified CM because Unified CVP is the primary routing point. This allows higher scalability per cluster compared to Unified IP IVR implementations. See Chapter 10, "Sizing Unified CCE Components and Servers"for more information.

Best Practices

- WAN connections to agent sites must be provisioned with bandwidth for voice as well as control and CTI. SeeBandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20 for more information.
- A local voice gateway might be needed at remote sites for local out-calling and 911.

Centralized Voice Gateways with Centralized Call Treatment and Queuing Using Unified System CCE 7.x with Unified CVP

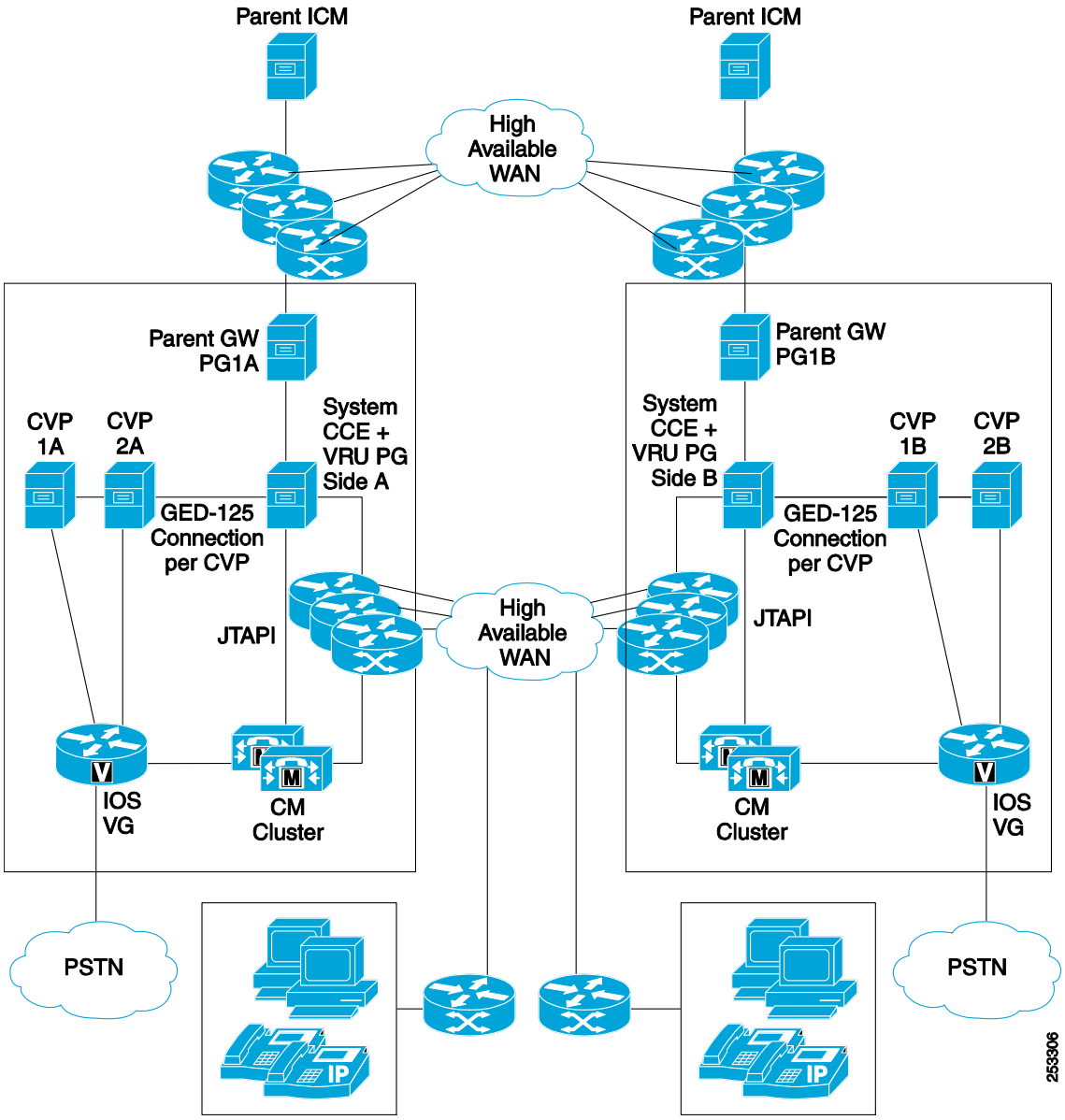
Load balancing of calls across Unified CVP Call Servers is managed by SIP and Cisco Unified SIP Proxy (CUSP). The load balancing does not take into account the site where the call came in, but calls are distributed based on simple load balancing rules define in Cisco Unified Presence (for example, alternate call distributions across configured Unified CVP Call Servers and preferential weighting of Call Servers).

Currently, if the system is designed to do so, Unified CVP can queue the call at the ingress gateway. This requires that Unified CVP be configured with **settransferlabel** for H.323 or **Send To Originator** for SIP, to match the NetworkVRU label. This will cause Unified CVP to send the call back to the ingress gateway for queuing when a label matching this NetworkVRU label is returned from Unified ICM/CCE. Currently Unified ICM/CCE is unaware of the location of the initial gateway, therefore it cannot make a label selection based on the original ingress location of the call.

Considerations for Clustering Over the WAN

Figure 2-22 illustrates a deployment with clustering over the WAN.

Figure 2-22 Clustering Over the WAN



The following guidelines and considerations apply to deployments with clustering over the WAN:

- The network deployment supports highly available, converged Visible and Private Networks. The Unified ICM and Unified CCE Central Controller's Private traffic and Visible (Public) traffic are isolated and converge on different edge devices.
- WAN considerations for communications between the two Data Centers may include a Multiprotocol Label Switching (MPLS) backbone with VPN routing and forwarding table VRFs.
- Design the network to prevent any single points of failure. The visible network and private network(s) need to converge on separate switches and routers before connecting to the WAN.
- Isolation of the private network is not required. Central Controllers and Unified CCE System PGs can share a common private network path.

- Multiple private network paths can be provisioned. (Central Controllers and Unified CCE System PGs have separate private networks.)
- Bandwidth must be guaranteed across the WAN for the private network path traffic and visible (public) network traffic, with appropriate traffic prioritization. For more information, refer to Chapter 12, "Bandwidth Provisioning and QoS Considerations".
- Currently there is no bandwidth calculator for the private network bandwidth between the gateway and system PG pairs because this has not been certified. For guidance, refer to the section onBandwidth Provisioning, page 12-16.
- A side-to-side private network of duplexed Central Controller and PGs has a maximum one-way latency of 100 ms, but 50 ms is preferred.

The underlying network infrastructure for LAN and WAN provisioning must meet all the above requirements. Key factors are isolation of visible and private paths as well as critical low-latency and bandwidth, especially on the private path. The isolated private networks for PGs and Central Controllers provide some degree of independence from each other's private link failures. The more path/route diversity provisioned, the greater the fault tolerance. For example, if the private network between the parent central controllers goes down, the child central controllers can still continue to function in duplex mode.

MPLS Considerations

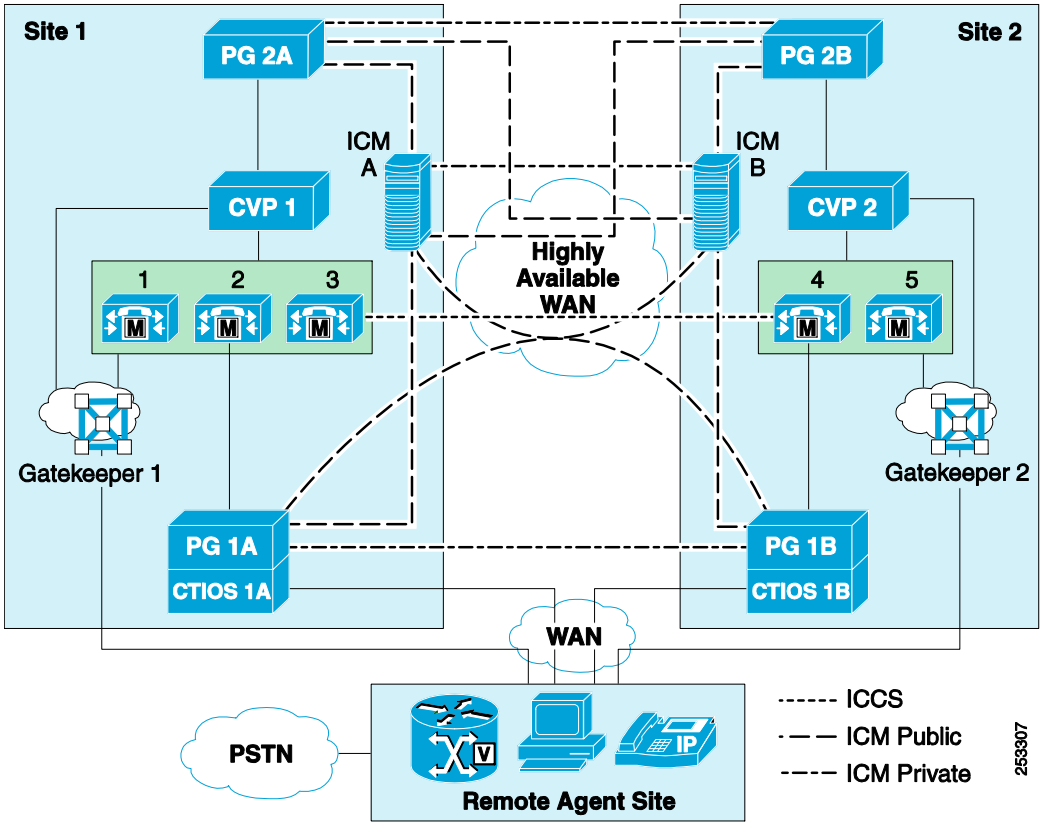
If an MPLS network can guarantee the route diversity, latency, and bandwidth, and if it is configured to support label switch paths that route via independent topologies and hardware elements to meet the above requirements, then the application will work as designed. It is important to ensure that the route autonomy is not compromised over time through adaptive change.

For additional information regarding best practices and high availability deployments, refer to the section on IPT: Clustering Over the WAN , page 2-32.

Distributed Voice Gateways with Distributed Call Treatment and Queuing Using Unified CVP

In this model, the voice gateways are VoiceXML gateways distributed to agent locations. Unified CVP is centrally located and used for treatment and queuing on the remote gateways. Figure 2-23 illustrates this model.

Figure 2-23 Distributed Voice Gateways with Distributed Call Treatment and Queuing Using Unified CVP



Advantages

- No or minimal voice RTP traffic across WAN links if ingress calls and gateways are provisioned to support primarily their local agents. Transfers and conferences to other sites would traverse the WAN.
- Calls are treated and queued at the agent site, eliminating the need for queuing across a WAN connection.
- Local calls incoming and outgoing, including 911, can share the local VoiceXML gateway.
- There is less load on Unified CM because Unified CVP is the primary routing point. This allows higher scalability per cluster compared to Unified IP IVR implementations. See Sizing Unified CCE Components and Servers, page 10-1, for more information.

Best Practices

- Distributed gateways require minimal additional remote maintenance and administration over centralized gateways.
- The media server for Unified CVP may be centrally located or located at the agent site. Media may also be run from gateway flash. Locating the media server at the agent site reduces bandwidth requirements but adds to the server count and maintenance costs due to an additional server.

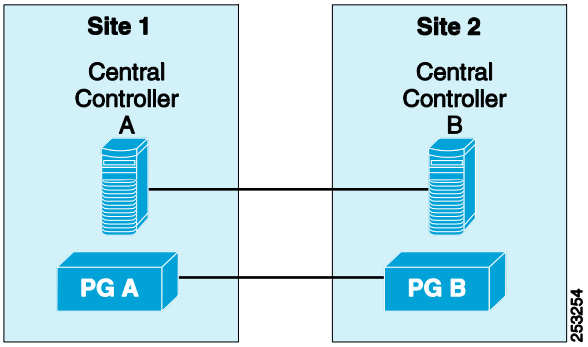
Site-to-Site Unified CCE Private Communications Options

Unified CCE private communications must travel on a separate path from the public communications between Unified CCE components. There are two options for achieving this path separation: dual and single links.

Unified CCE Central Controller Private and Unified CM PG Private Across Dual Links

Dual links, shown in Figure 2-24, separate Unified CCE Central Controller Private traffic from VRU/CM PG Private traffic.

Figure 2-24 Unified CCE Central Controller Private and Unified CM PG Private Across Dual Links



Advantages

- Failure of one link does not cause both the Unified CCE Central Controller and PG to enter simplex mode, thus reducing the possibility of an outage due to a double failure.
- The QoS configuration is limited to two classifications across each link, therefore links are simpler to configure and maintain.
- Resizing or alterations of the deployment model and call flow might affect only one link, thus reducing the QoS and sizing changes needed to ensure proper functionality.
- Unanticipated changes to the call flow or configuration (including misconfiguration) are less likely to cause issues across separate private links.

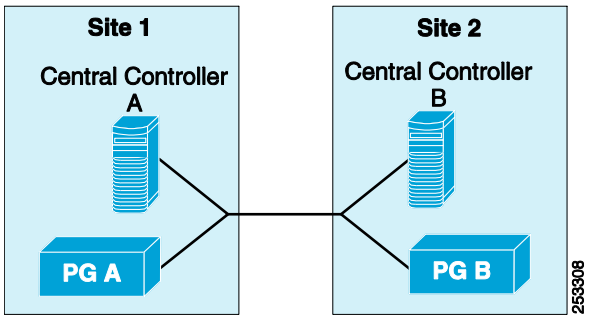
Best Practices

- Deploy separate links across separate dedicated circuits. The links, however, do not have to be redundant and must not fail-over to each other.
- Link sizing and configuration must be examined before any major change to call load, call flow, or deployment configuration.
- Make the link a dedicated circuit and not tunneled across the highly available (HA) WAN. See Best Practices, at the beginning of the section on IPT: Clustering Over the WAN , page 2-32, for more information on path diversity.

Unified CCE Central Controller Private and Unified CM PG Private Across Single Link

A single link, shown in Figure 2-25, carries both Unified CCE Central Controller Private traffic and VRU/CM PG Private traffic. Single-link implementations are more common and less costly than dual-link implementations.

Figure 2-25 Unified CCE Central Controller Private and Unified CM PG Private Across a Single Link



Advantages

- Less costly than separate-link model.
- Fewer links to maintain, but more complex.

Best Practices

- The link does not have to be redundant. If a redundant link is used, however, latency on failover must not exceed 500 ms.
- Separate QoS classifications and reserved bandwidth are required for Central Controller high-priority and PG high-priority communications. For details, see Bandwidth Provisioning and QoS Considerations, page 12-1.
- Link sizing and configuration must be examined before any major change to call load, call flow, or deployment configuration. This is especially important in the single-link model.
- Make the link a dedicated circuit fully isolated from, and not tunneled across, the highly available (HA) WAN. See Best Practices, at the beginning of the section on IPT: Clustering Over the WAN , page 2-32, for more information on path diversity.

Failure Analysis of Unified CCE Clustering Over the WAN

This section describes the behavior of clustering over the WAN for Unified CCE during certain failure situations. The stability of the highly available (HA) WAN is extremely critical in this deployment model, and failure of the highly available WAN is considered outside the bounds of what would normally happen.

For illustrations of the deployment models described in this section, refer to the figures shown previously for IPT: Clustering Over the WAN , page 2-32.

Entire Central Site Loss

Loss of the entire central site is defined as the loss of all communications with a central site, as if the site were switched off. This can result from natural disasters, power issues, major connectivity issues, and human error, among other things. If a central site retains some but not all connectivity, it is not considered a site loss but rather a partial connectivity loss, and this scenario is covered in subsequent sections.

When an entire central site has completely lost Unified CCE clustering over the WAN, Remote Agents will fail-over properly to the redundant site. Failover times can range from 1 to 60 seconds for agents. Variations are due to agent count, phone registration location, and agent desktop server used.

When using distributed VoiceXML gateways and Unified CVP, the gateways must fail-over from one site to another if their primary site is lost. This failover takes approximately 30 seconds, and calls coming into the remote gateways during those 30 seconds will be lost.

Private Connection Between Site 1 and Site 2

If the private connection between Unified CCE Central Controller sides A and B fails, one Unified CCE Call Router will go out-of-service and the other Unified CCE Call Router will then be running in simplex mode until the link is reinstated. PGs with active connections to the Unified CCE Router, that goes out-of-service, realign message streams to the remaining active Unified CCE Router side. This situation will not cause any call loss or failure.

If the private connection between PG side A and PG side B fails, the disabled synchronizer (side B) initiates a test of its peer synchronizer via the TOS procedure on the Public or Visible Network connection. If PG side B receives a TOS response stating that the A side synchronizer is enabled or active, then the B side immediately goes out of service, leaving the A side to run in simplex mode until the Private Network connection is restored. The PIM, OPC, and CTI SVR processes become active on PG side A, if not already in that state, and the CTI OS Server process still remains active on both sides as long as the PG side B server is healthy. If the B side does not receive a message stating that the A side is enabled, then side B continues to run in simplex mode and the PIM, OPC, and CTI SVR processes become active on PG side B if not already in that state. This condition (PG side B going Active) occurs only if the PG side A server is truly down or unreachable due to a double failure of visible and private network paths. The side A or Side B PG runs in simplex mode until the private link is reinstated.

There is no impact to the agents, calls in progress, or calls in queue because the agents stay connected to their already established CTI OS Server process connection. The system can continue to function normally; however, the PGs will be in simplex mode until the private network link is restored.

When using a combined private link, Unified CCE Central Controller and PG private connections will be lost if the link is lost. This will cause both components to switch to simplex mode as described above. This situation will not cause any call loss or failure.

Connectivity to Central Site from Remote Agent Site

If connectivity to one of the central sites is lost from a Remote Agent site, all phones and agent desktops will immediately switch to the second central site and begin processing calls. Failover typically takes between 1 and 60 seconds.

Highly Available WAN Failure

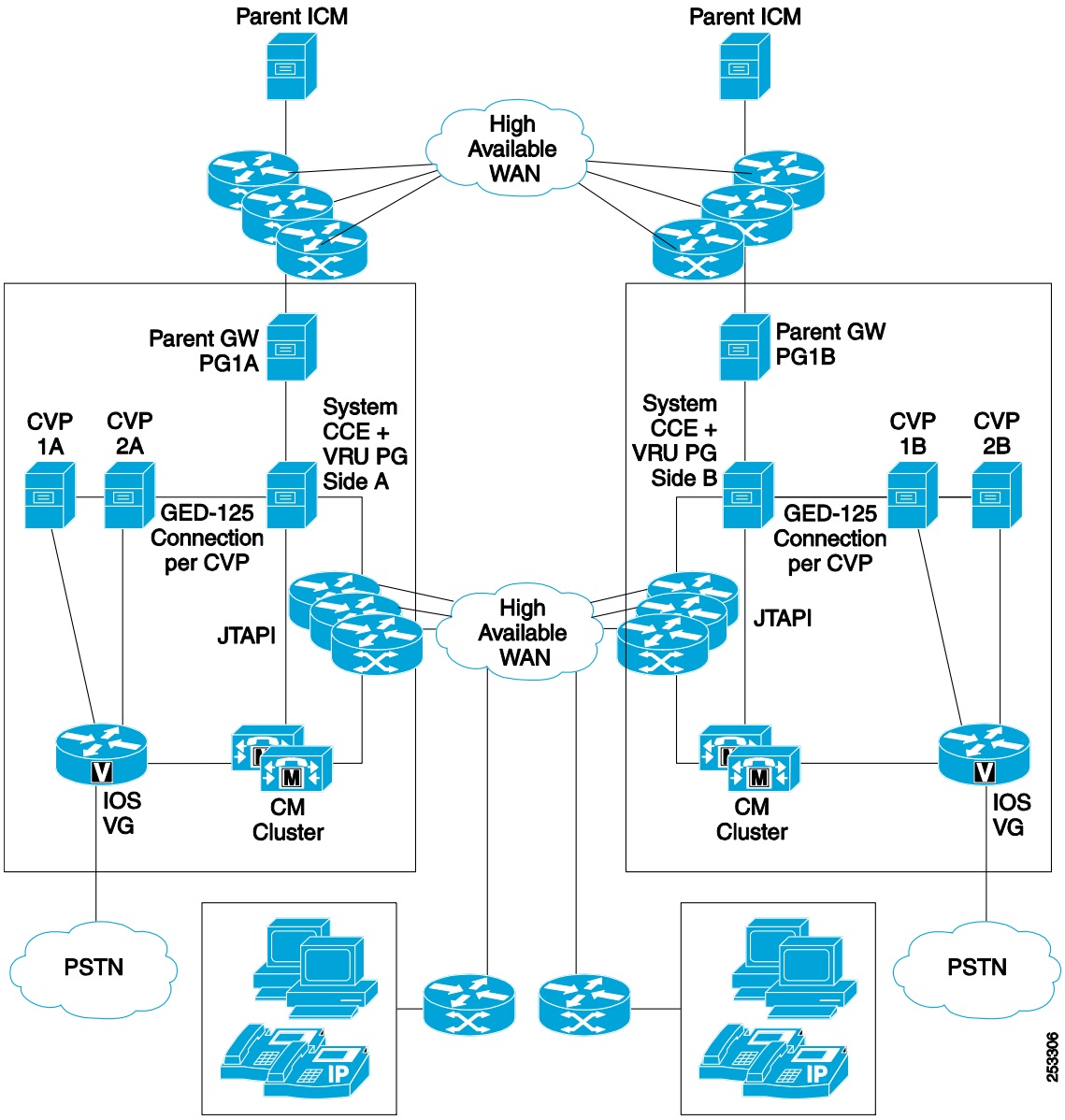
By definition, a highly available (HA) WAN does not fail under normal circumstances. If the HA WAN is dual-path and fully redundant, a failure of this type would be highly unusual. This section discusses what happens in this unlikely scenario.

If the HA WAN is lost for any reason, the Unified CM cluster becomes split. The primary result from this occurrence is that Unified CCE loses contact with half of the agent phones. Unified CCE is in communication with only half of the cluster and cannot communicate with or see any phones registered on the other half. This causes Unified CCE to immediately log out all agents with phones that are no longer visible. These agents cannot log back in until the highly available WAN is restored or their phones are forced to switch cluster sides.

Split Unified CCE Gateway PGs

To enhance the distributed architecture of the Unified System CCE 7.x deployment, the support for geographically distributed Cisco Unified CCE Gateway PGs is needed. The Unified CCE Gateway PGs are deployed in the same location as the System PGs, adding maximum recovery capabilities in the event of a site failure. Figure 2-26 shows a distributed Unified System CCE 7.x deployment supporting two Remote Data Centers with Unified CCE Gateway PGs co-located (on separate servers) with each of the distributed Unified System CCE7.x systems. Note that the same would apply if the children were Unified CCEs utilizing the Unified CCE System PG.

Figure 2-26 Gateway PG Co-Located



Remote Agent Over Broadband

An organization might want to deploy Unified CCE to support remote agents (for example, at-home agents) using a Cisco Unified IP Phone over a broadband internet connection. This section outlines the remote agent solution that can be deployed using a desktop broadband asymmetric digital subscriber line (ADSL) or Cable connection as the remote network. Another option is to use the Cisco Unified Mobile Agent solution (for details, see Cisco Unified Mobile Agent , page 6-1). Both Cisco Unified Mobile Agent and Remote Agent over Broadband can be supported concurrently using the same back-end infrastructure with the Cisco Virtual Office solution, which is an underlying end-to-end secure infrastructure for remote teleworkers utilizing a converged VPN architecture.

The Cisco Voice and Video Enabled IPSec VPN (V3PN) ADSL or Cable connection can use a Cisco 800 Series router as an edge router to the broadband network. The Cisco 800 Series router can provide the remote agent with V3PN, Encryption, Network Address Translation (NAT), Firewall, Cisco IOS Intrusion Detection System (IDS), and QoS on the broadband network link to the Unified CCE campus. Remote agent V3PN aggregation on the campus is provided via LAN to LAN VPN routers.

Use the Cisco 800 Series router with the following features for remote agent over broadband:

- Quality of Service (QoS) with Low-Latency Queuing (LLQ) and Class-Based Weighted Fair Queuing (CBWFQ) support
- Managed Switch
- Power over Ethernet (optional)

The Cisco 830, 870, and 880 Series routers are examples of acceptable routers. Avoid using the Cisco 850 and 860 Series routers for this application because they have limited QoS feature support.

Advantages

- A remote-agent deployment results in money saved for a contact center enterprise, thereby increasing return on investment (ROI).
- Remote agents can be deployed with standard Unified CCE agent desktop applications such as Cisco CTI OS, Cisco Agent Desktop, or customer relationship management (CRM) desktops.
- The Broadband Agent Desktop Always-on connection is a secure extension of the corporate LAN in the home office.
- Remote agents have access to the same Unified CCE applications and most Unified CCE features in their home office as when they are working at the Unified CCE contact center, and they can access those features in exactly the same way.
- The remote-agent router can provide high-quality voice using IP phones, with simultaneous data to the agent desktop via existing broadband service.
- Unified CCE home agents and home family users can securely share broadband Cable and DSL connections, with authentication of Unified CCE corporate users providing access to the VPN tunnel.
- The remote-agent routers can be managed centrally by the enterprise using a highly scalable and flexible management product such as Cisco Unified Operations Manager.
- The remote-agent-over-broadband solution is based on Cisco IOS VPN Routers for resiliency, high availability, and a building-block approach to high scalability that can support thousands of home agents.
- All traffic, including data and voice, is encrypted with the Triple Data Encryption Standard (3DES).
- • The remote-agent router can be deployed as part of an existing Unified CM installation.
- Remote agents can have the same extension type as campus agents.

Best Practices

- Follow all applicable V3PN and Cisco Virtual Office design guidelines outlined in the documentation available at:
  - http://www.cisco.com/go/cvo
  - http://www.cisco.com/go/designzone
- Configure remote agent IP phones to use G.729 with minimum bandwidth limits. Higher-quality voice can be achieved with the G.711 codec. The minimum bandwidth to support G.711 is 512 kbps upload speed.
- Implement fault and performance management tools such as NetFlow, Service Assurance Agent (SAA), and Internetwork Performance Monitor (IPM).
- Wireless access points are supported; however, their use is determined by the enterprise security polices for Remote Agents.
- Only one remote agent per household is supported.
- Configure the conference bridge on a DSP hardware device. There is no loss of conference voice quality using a DSP conference bridge. This is the preferred solution even for pure Cisco Unified Communications deployments.
- The remote-agent-over-broadband solution is supported only with centralized Unified CCE and Unified CM clusters.
- There might be times when the ADSL or Cable link goes down. When the link is back up, you might have to reset your ADSL or Cable modem remote agent router and IP phone. This task will require Remote Agent training.
- Only unicast Music on Hold (MoH) streams are supported.
- There must be a Domain Name System (DNS) entry for the remote agent desktop, otherwise the agent will not be able to connect to a CTI server. DNS entries can be updated dynamically or entered as static updates.
- The remote agent workstation and IP phone must be set up to use Dynamic Host Configuration Protocol (DHCP).
- The remote agent workstation requires Windows XP Pro for the operating system. In addition, XP Remote Desktop Control must be installed.
- The Cisco Unified IP Phone requires a power supply if the remote-agent router does not have the Power over Ethernet option.
- Remote agent broadband bandwidth requires a minimum of 256 kbps upload speed and 1.4 Mbps download speed for ADSL, and 1 Mbps download for Cable. Before actual deployment, make sure that the bandwidth is correct. If you are deploying Cable, then take into account peak usage times. If link speeds fall below the specified bandwidth, the home agent can encounter voice quality problems such as clipping.
- Remote agent round-trip delay to the Unified CCE campus is not to exceed 180 ms for ADSL or 60 ms for Cable. Longer delay times can result in voice jitter, conference bridge problems, and delayed agent desktop screen pops.
- If the Music on Hold (MoH) server is not set up to stream using a G.729 codec, then a transcoder must be set up to enable outside callers to receive MoH.

- For Cisco Supervisor Desktop, there are supervisor limitations to silent monitoring, barge-in, intercept, and voice recording with regard to home agent IP phones. Cisco Agent Desktop (Enterprise and Express) home and campus supervisors cannot voice-monitor home agents. Supervisors are capable of sending and receiving only text messages, and they can see which home agents are online and can log them out.
- CTI OS Supervisor home and campus supervisors can silently monitor, barge in, and intercept, but not record home agents. CTI OS home and campus supervisors can send and receive text messages, make an agent ready, and also log out home agents.
- Connect the agent desktop to the RJ45 port on the back of the IP phone. Otherwise, CTI OS Supervisor will not be able to voice-monitor the agent phone.
- Only IP phones that are compatible with Cisco Unified CCE are supported. For compatibility information, refer to the following documentation:
  - *Hardware and System Software Specification (Bill of Materials) for Cisco ICM/CCE Enterprise & Hosted Editions, Release 8.0(1)*, available at

    http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_implementation_design_guides_list.html
  - *IPCC Enterprise Software Compatibility Guide*, available at

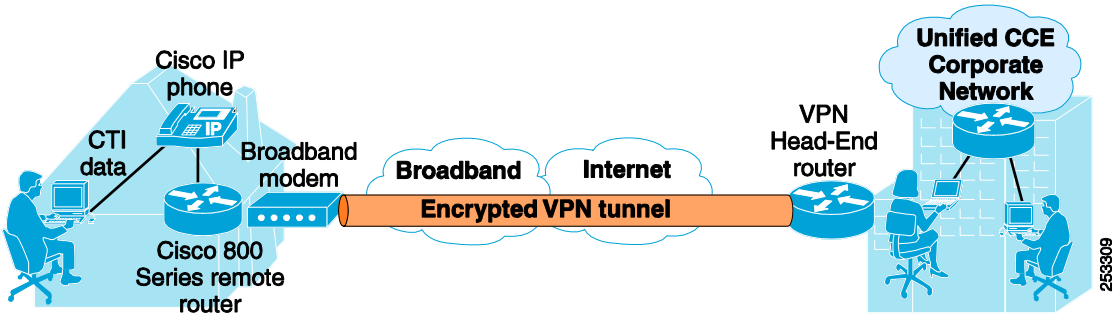
    http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_device_support_tables_list.html
  - *Release Notes for Cisco Unified ICM/Contact Center Enterprise & Hosted, Release 8.0(1)*, available at

    http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_release_notes_list.html
- You can find a test for the broadband line speed at http://www.Broadbandreports.com. From this website, you can execute a test that will benchmark the home agent's line speed (both upload and download) from a test server.

Remote Agent with Unified IP Phones Deployed via the Cisco Virtual Office Solution

In this model, the Remote Agent's IP phone and workstation are connected via the VPN tunnel to the main Unified CCE campus. Customer calls routed to the remote agent are handled in the same manner as campus agents. (See Figure 2-27.)

Figure 2-27 Remote Agent with IP Phones Deployed via the Cisco Virtual Office Solution



Advantages

- High-speed broadband enables cost-effective office applications.
- Site-to-site always-on VPN connection.

- Advanced security functions allow extension of the corporate LAN to the home office.
- Supports full range of converged desktop applications, including CTI data and high-quality voice.

Best Practices

- Minimum broadband speed supported is 256 kbps upload and 1.0 Mbps download for cable.
- Minimum broadband speed supported is 256 kbps upload and 1.4 Mbps download for ADSL.
- Agent workstation must have 500 MHz, 512 MB RAM or greater.
- IP phone must be configured to use G.729 on minimum broadband speeds.
- QoS is enabled only at the remote-agent router edge. Currently, service providers are not providing QoS.
- Enable security features on the remote-agent router.
- The Cisco 7200 VXR and Catalyst 6500 IPSec VPN Services Module (VPNSM) offer the best LAN-to-LAN performance for agents.
- The remote agent's home phone must be used for 911 calls.
- Use Redirect-on-no-answer (RONA) when a remote agent is logged in and ready but is unavailable to pick up a call.

Traditional ACD Integration

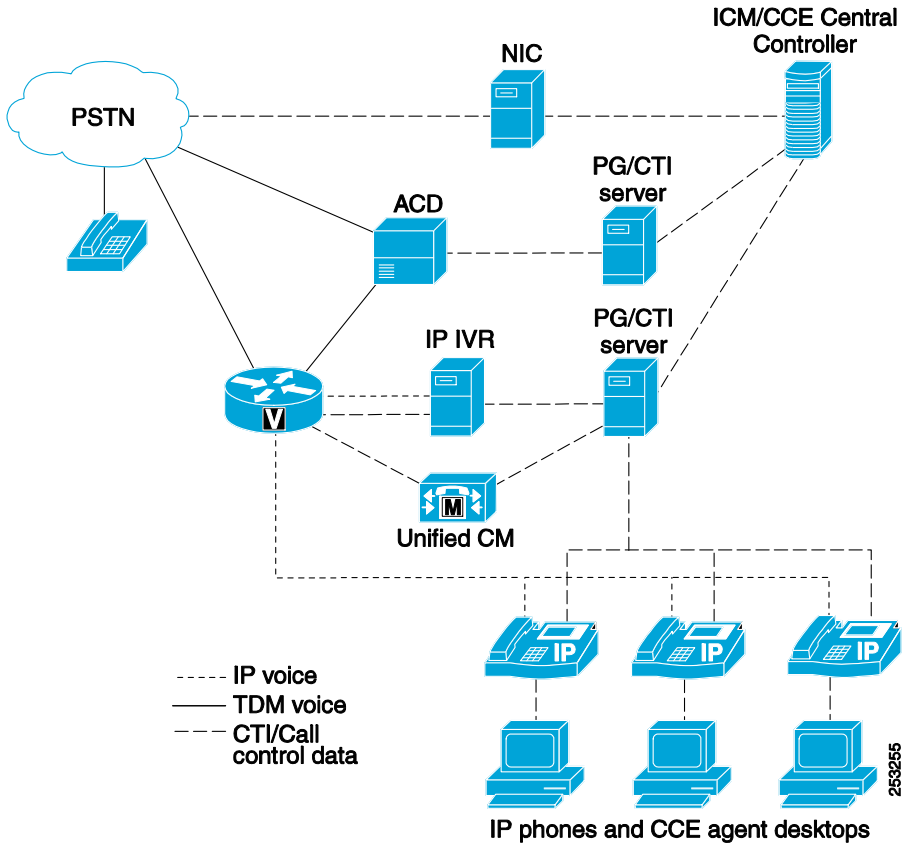
Enterprises that want to integrate traditional ACDs with their Unified CCE could use a parent/child deployment where the Unified ICM and Unified CCE each have a Central Controller, or a hybrid deployment where Unified ICM and Unified CCE use a shared Central Controller. Several options exist within those categories, depending on how the calls are routed within the deployment.

Hybrid Deployment with PSTN Prerouting

Enterprises that want to load-balance calls between a traditional ACD site and a Unified CCE site could add a pre-routing Network Interface Controller (NIC). (See Figure 2-28.) This requires that the Unified ICM have a NIC that supports the PSTN service provider. In this scenario, the PSTN will query the Unified ICM/CCE Central Controller (via the NIC) to determine which site is best, and the Unified ICM/CCE response back to the PSTN will instruct the PSTN where (which site) to deliver the call. Any call data provided by the PSTN to the Unified ICM/CCE will be passed to the agent desktop (traditional ACD or Unified CCE).

In order to transfer calls between the two sites (ACD site and Unified CCE site), a PSTN transfer service could be used. Use of a PSTN transfer service avoids any double trunking of calls at either site. An alternative to using a PSTN transfer service is to deploy TDM voice circuits between the traditional ACD and Unified CCE voice gateways. In that environment, any transfer of a call back to the original site will result in double trunking between the two sites. Each additional transfer between sites will result in an additional TDM voice circuit being utilized.

Figure 2-28 Integrating a Traditional ACD with a Unified CCE Site Using a Hybrid Deployment and Prerouting via PSTN



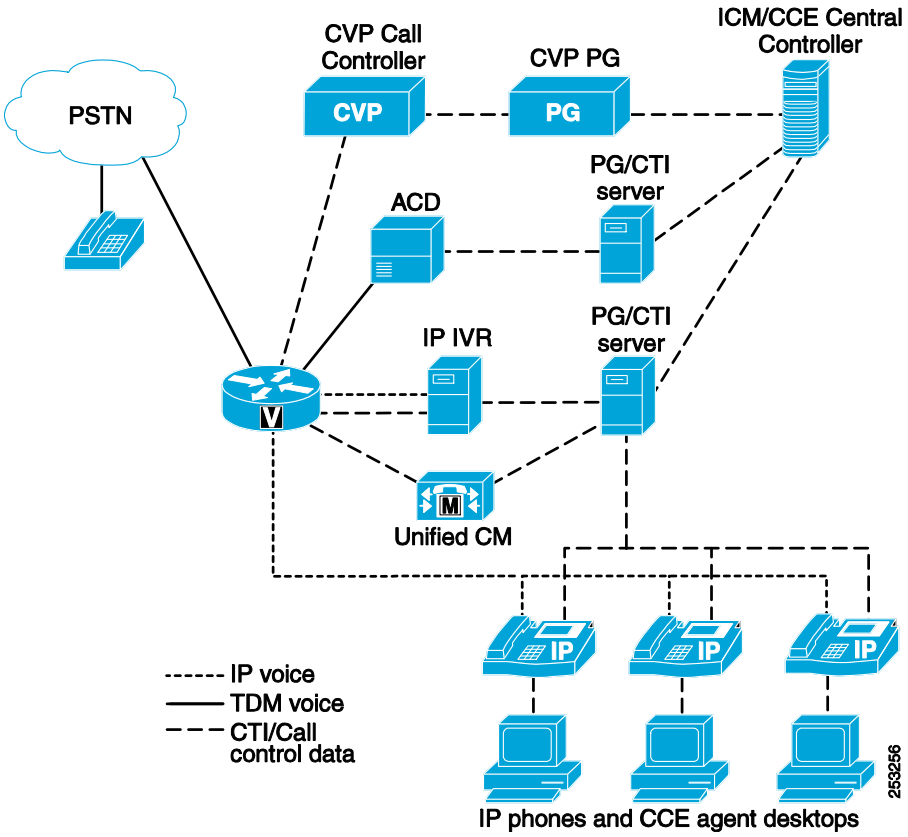
Hybrid Deployment with Fixed PSTN Delivery

An alternative to pre-routing calls from the PSTN is to have the PSTN deliver calls to just one site or to split the calls across the two sites according to some set of static rules provisioned in the PSTN. When the call arrives at either site, either the traditional ACD or the Unified CM will generate a route request to the hybrid Unified ICM/CCE to determine which site is best for this call. If the call needs to be delivered to an agent at the opposite site from where the call was originally routed, then TDM circuits between sites will be required. Determination of where calls are routed, and if and when they are transferred between sites, will depend upon the enterprise business environment, objectives, and cost components.

Hybrid Deployment with Unified CVP

Alternatively, customers may choose to front-end all calls with Unified CVP to provide initial call treatment and queuing across both the TDM ACD and Unified CCE agents. (See Figure 2-29.)

Figure 2-29 Integrating Unified CVP with a Traditional ACD and a Unified CCE Site Using a Hybrid Deployment and Unified CVP

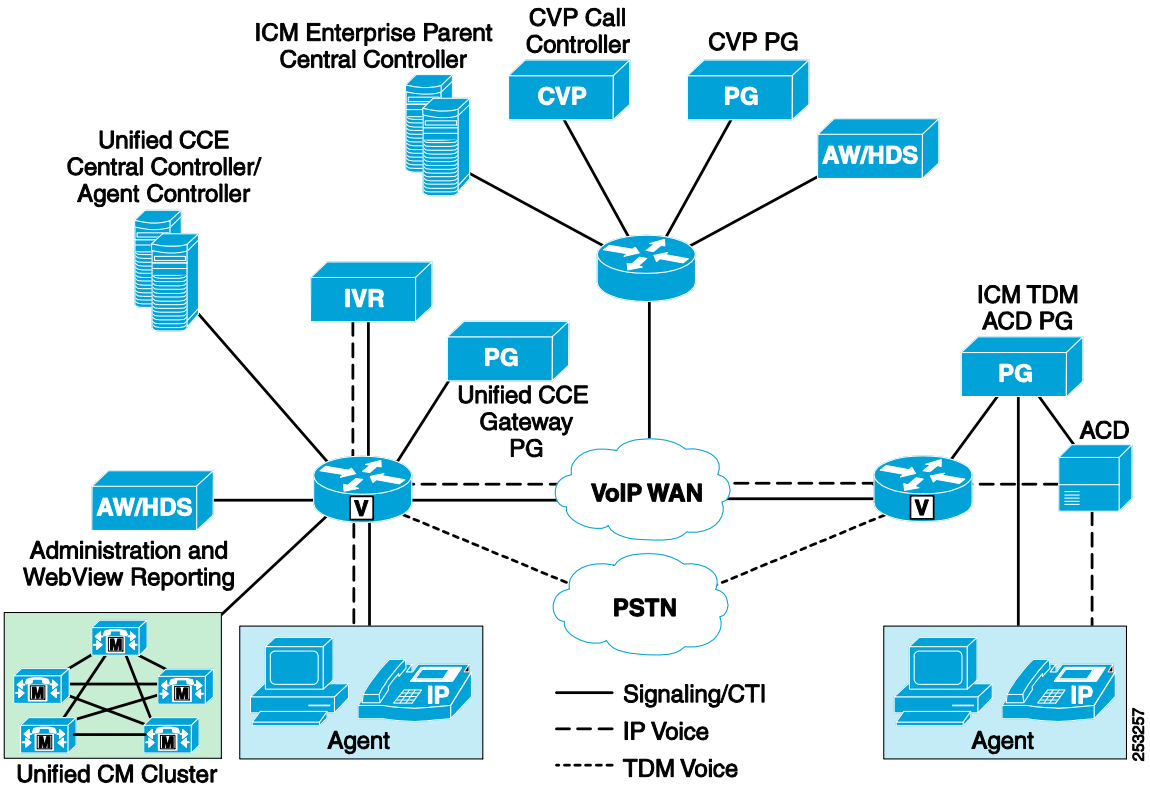


In this design, all calls first come to the voice gateway controlled by Unified CVP, and they are then directed by the Unified ICM/CCE Call Router. Unified ICM/CCE uses the PG connections to the TDM ACD and Unified CCE PG to monitor for available agents. Calls are queued in Unified CVP until an agent becomes available in either environment. When a call needs to be transferred to the TDM ACD, it will hairpin in the voice gateway, meaning that it comes into the gateway on a T1 interface from the PSTN carrier network and goes out on a second physical T1 interface to appear as a trunk on the TDM ACD. Most TDM ACDs are unable to accept inbound calls in IP from the voice gateway and require this physical T1 interface/connection. Unified CCE agents will receive their calls directly over the IP voice network.

Parent/Child Deployment

The parent/child model is illustrated in Figure 2-30.

Figure 2-30 Parent/Child Model for Integrating a Traditional ACD with a Unified CCE Site



In this model, the Unified ICM Enterprise parent has PGs connected to a Unified CCE (using System PG) at one site with a complete installation, and a second site with a TDM ACD that is using a Unified ICM TDM ACD PG. In this model, Unified ICM still provides virtual enterprise-wide routing, call treatment, and queuing with the distributed Unified CVP voice gateways at the sites. Unified ICM also has full visibility to all the sites for agents and calls in progress. The difference in this model is that Unified CCE provides local survivability. If it loses connection to the Unified ICM parent, the calls will still be treated locally just as they would be at the TDM ACD site.

Traditional IVR Integration

There are numerous ways that traditional IVRs can be integrated into a Unified CCE deployment. Determination of which way is best will depend upon many factors that are discussed in the following sections. The primary consideration, though, is determining how to eliminate or reduce IVR double trunking when transferring the call from the IVR.

Using PBX Transfer

Many call centers have existing traditional IVR applications that they are not prepared to rewrite. In order to preserve these IVR applications, but yet integrate them into a Unified CCE environment, the IVR must have an interface to Unified CCE. (See Figure 2-31.)

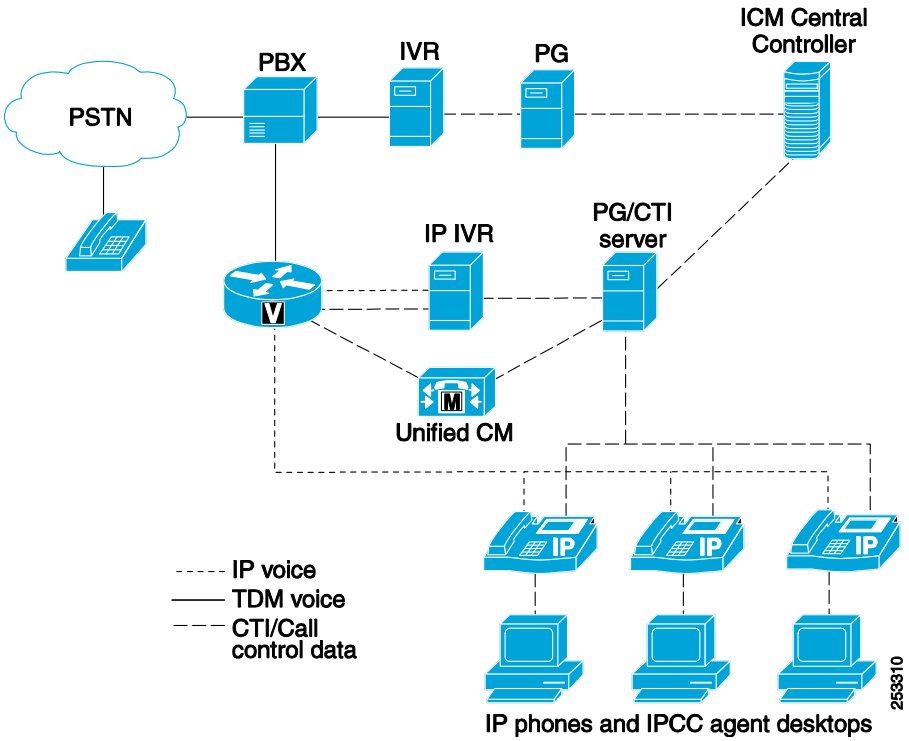
There are two versions of the IVR interface to Unified CCE. One is simply a post-routing interface (Call Routing Interface, or CRI), which just allows the IVR to send a post-route request with call data to Unified CCE. Unified CCE returns a route response instructing the IVR to transfer the call elsewhere.

In this scenario, the traditional IVR will invoke a PBX transfer to release its port and transfer the call into the Unified CCE environment. Any call data passed from the IVR will be passed by Unified CCE to the agent desktop or Unified IP IVR.

The other IVR interface to Unified CCE is the Service Control Interface (SCI). The SCI allows the IVR to receive queuing instructions from Unified CCE. In the PBX model, the SCI is not required.

Even if the IVR has the SCI interface, it is still preferable to deploy Unified CVP or Unified IP IVR for all call queuing because this prevents any additional utilization of the traditional IVR ports. In addition, use of the Unified IP IVR for queuing provides a way to requeue calls on subsequent transfers or RONA treatment.

Figure 2-31 Traditional IVR Integration Using PBX Transfer



In this design, calls come first to the PBX from the PSTN carrier network on a standard T1 trunk interface. The PBX typically uses a hunt group to transfer the call to the IVR, putting all of the IVR ports into the hunt group as agents in auto available mode. The PBX looks like the PSTN to Unified CCE because it does not have a PG connected to the PBX. Unified CCE cannot track the call from the original delivery to the IVR, and it will have reporting only from the time the call arrived at the IVR and the IVR informed Unified CCE of the call.

When the caller opts out of the IVR application, the IVR sends a Post-Route to Unified CCE using the Call Routing Interface (CRI). Because this application does not require calls to queue in the IVR, the CRI would be the preferred interface option. Unified CCE will look at the agent states across the system and select the agent to send the call to (via agent phone number or device target) or translation-route the call to the Unified IP IVR for queuing.

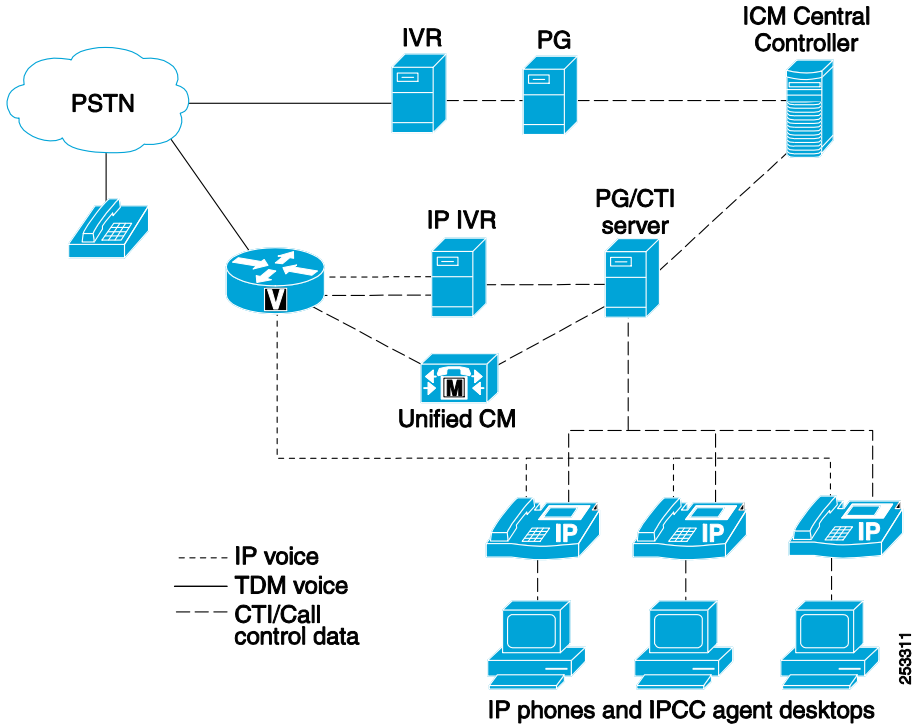
When the call is sent to an agent or into queue, it is hairpinned in the PBX, coming in from the PSTN on a T1 trunk port and then going out to a voice gateway on a second T1 trunk port in the PBX. This connection is used for the life of the call.

Alternatively, if you want to track the call from its entry at the PBX or if you need to capture the caller ANI or original dialed number, you can install a PG on the PBX. The PBX can request (via a Post-Route to Unified CCE) which IVR port to send the call to behind the PBX. The PBX cannot use a hunt group to deliver the call from the PBX to the IVR. Unified CCE requires direct DNIS termination to ensure that the translation route maintains the call data collected in the PBX and makes it available to the IVR.

Using PSTN Transfer

This model is very similar to the previous model, except that the IVR invokes a PSTN transfer (instead of a PBX transfer) so that the traditional IVR port can be released. (See Figure 2-32.) Again, the Unified IP IVR would be used for all queuing so that any additional occupancy of the traditional IVR ports is not required and also so that any double trunking in the IVR is avoided. Any call data collected by the traditional IVR application will be passed by Unified CCE to the agent desktop or Unified IP IVR.

Figure 2-32 Traditional IVR Integration Using PSTN Transfer



In this model, the TDM IVR is set up as a farm of IVR platforms that have direct PSTN connections for inbound calls. The IVR has a PG connection to Unified CCE, which tracks all calls in the system. When a caller opts out of the IVR treatment, the IVR sends a post-route request to Unified CCE, which returns a label that will direct the call either to an agent or to the Unified IP IVR for queuing.

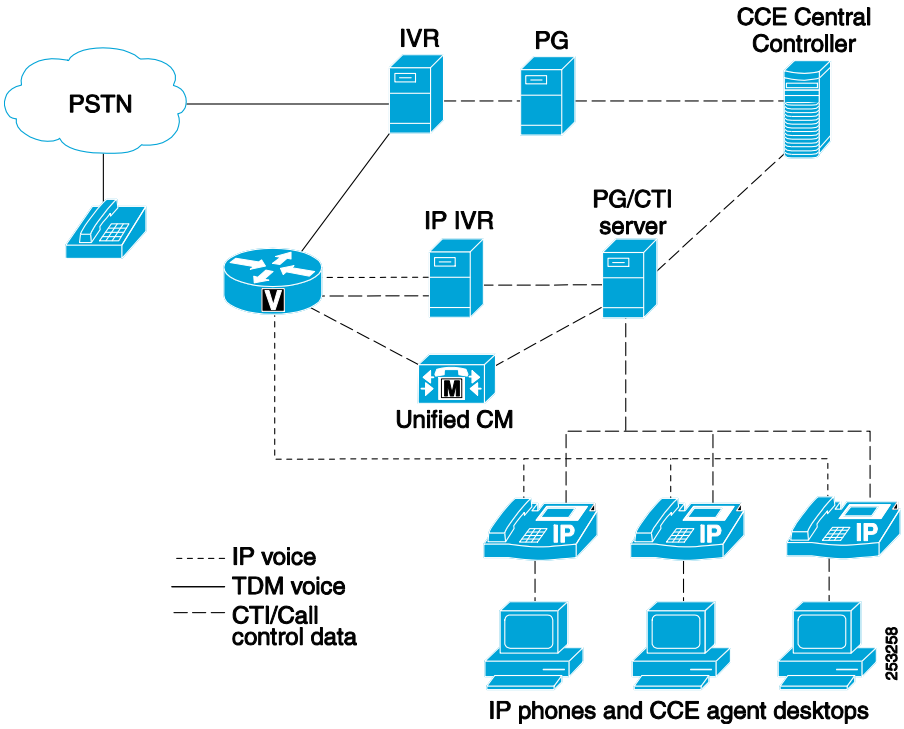
The label that is returned to the TDM IVR instructs it to send an in-band transfer command using transfer tones (*8 with a destination label in the carrier network). The IVR has to outpulse these tones to the service provider with tone generation or play the tones via a recorded file.

Using IVR Double Trunking

If your traditional IVR application has a very high success rate, where most callers are completely self-served in the traditional IVR and only a very small percentage of callers ever need to be transferred to an agent, then it might be acceptable to double-trunk the calls in the traditional IVR for that small

percentage of calls. (See Figure 2-33.) Unlike the previous model, if the traditional IVR has a Service Control Interface (SCI), then the initial call queuing could be done on the traditional IVR. The reason this is beneficial is that, in order to queue the call on the Unified IP IVR, a second traditional IVR port would be used to transfer the call to the Unified IP IVR. By performing the initial queuing on the traditional IVR, only one traditional IVR port is used during the initial queuing of the call. However, any subsequent queuing as a result of transfers or RONA treatment must be done on the Unified IP IVR to avoid any double trunking. If the traditional IVR does not have an SCI interface, then the IVR will just generate a post-route request to Unified CCE to determine where the call is transferred. All queuing in that scenario would have to be done on the Unified IP IVR.

Figure 2-33 Traditional IVR Integration Using IVR Double Trunking



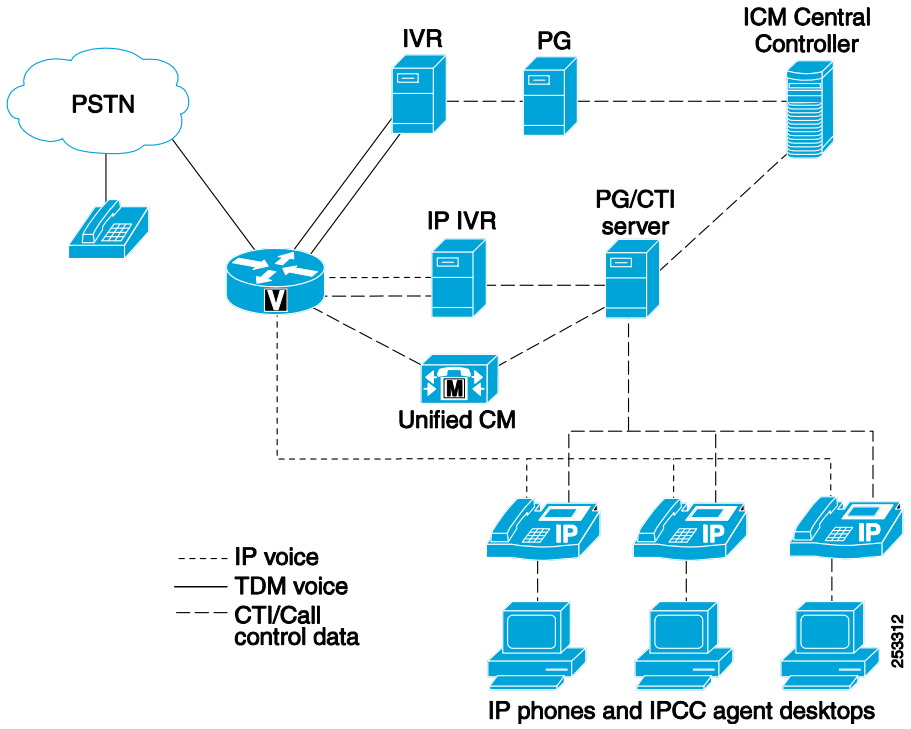
In this model, the TDM IVR is set up as a farm of IVR platforms that have direct PSTN connections for inbound calls. The IVR has a PG connection to Unified CCE, which tracks all calls in the system. When a caller opts out of the IVR treatment, the IVR sends a post-route request to Unified CCE, which returns a label that will either direct the call to an agent or queue the call locally on the TDM IVR using the Service Control Interface (SCI). The transfer to the agent is done by the TDM IVR selecting a second port to hairpin the call to the voice gateway and to the Unified CCE agent. This takes up two ports for the time the call is at the agent.

Using Unified CM Transfer and IVR Double Trunking

Over time, it might become desirable to migrate the traditional IVR applications to Unified CVP or Unified IP IVR. However, if a small percentage of traditional IVR applications still exist for very specific scenarios, then the IVR could be connected to a second voice gateway. (See Figure 2-34.) Calls arriving at the voice gateway from the PSTN would be routed by Unified CM. Unified CM could route specific DNs to the traditional IVR or let Unified CCE, Unified CVP, or Unified IP IVR determine when to transfer calls to the traditional IVR. If calls in the traditional IVR need to be transferred to a Unified

CCE agent, then a second IVR port, trunk, and voice gateway port would be used for the duration of the call. Ensure that transfer scenarios do not allow multiple loops to be created because voice quality could suffer.

Figure 2-34 Traditional IVR Integration Using Unified CM Transfer and IVR Double Trunking

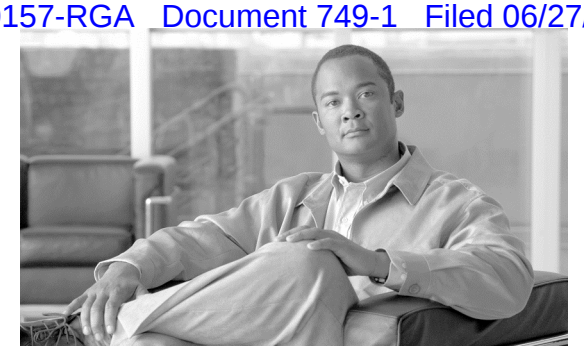


In this model, the TDM IVR is front-ended by either Unified CVP using the voice gateway or the Unified IP IVR and Unified CM with Unified CCE to determine the location to provide call treatment.

With Unified CVP, calls coming into the voice gateway would immediately start a routing dialog with Unified CCE using the Service Control Interface (SCI). Based upon the initial dialed number or prompting in Unified CVP, Unified CCE would decide if the call needs to be sent to the TDM IVR for a specific self-service application or if Unified CVP has the application available for the caller. If the call was sent to the TDM IVR, the TDM IVR sends a route request to Unified CCE when the caller opts out. The reply is not sent back to the TDM IVR but back to Unified CVP as the original routing client. Unified CVP would then take the call leg away from the TDM IVR and transfer it to the Unified CCE agent over the VoIP network or hold it in queue locally in the voice gateway.

With Unified CM, calls coming into the voice gateway would hit a CTI route point for Unified CM to send a route request to Unified CCE to determine the appropriate call treatment device for the caller. If the CTI route point indicated an application that still is on the TDM IVR, Unified CCE would instruct Unified CM to transfer the call to the TDM IVR by hairpinning the call using a second T1 port on the voice gateway to connect to the TDM IVR. Unified CCE could also instruct Unified CM to translation-route the call to the Unified IP IVR for call processing or prompting, then make a subsequent transfer to the TDM IVR for further processing. When the caller opts out of the TDM IVR, it sends a post-route request to Unified CCE, and Unified CCE returns a label to the TDM IVR. This label instructs the TDM IVR to transfer the call using a second T1 port on the IVR and to pass the call back to the voice gateway and over to the Unified CCE agent under Unified CM's dial plan.

In the model controlled by Unified CM, calls are initially received by the voice gateway and hairpinned to the TDM IVR on a second T1 port. When the IVR sends the call back to the Unified CCE agent, it uses a second TDM IVR port and a third port on the voice gateway. All three ports would be tied up on the voice gateway as long as the agent is talking with the caller, and both of the TDM IVR ports would be tied up for the duration of this call as well.



CHAPTER 3

Design Considerations for High Availability

Note Many of the design considerations and illustrations throughout this chapter have been revised and updated. Review the entire chapter before designing a Unified CCE system.

Designing for High Availability

Cisco Unified CCE is a distributed solution that uses numerous hardware and software components, and it is important to design each system in a way that eliminates any single point of failure or that at least addresses potential failures in a way that will impact the fewest resources in the contact center. The type and number of resources impacted will depend on how stringent your requirements are, the budget for fault tolerance, and which design characteristics you choose for the various Unified CCE components, including the network infrastructure. A good Unified CCE design will be tolerant of most failures (defined later in this section), but not all failures can be made transparent.

Cisco Unified CCE is a solution designed for mission-critical contact centers. The successful design of any Unified CCE deployment requires a team with experience in data and voice internetworking, system administration, and Unified CCE application design and configuration.

Note Simplex deployments are allowed for demo, laboratory, and non-production deployments. However, all production deployments *must* be deployed with redundancy for the core Unified CCE components (Call Routers, Loggers, PGs, and pre-routing gateways).

Before implementing Unified CCE, use careful preparation and design planning to avoid costly upgrades or maintenance later in the deployment cycle. Always design for the worst possible failure scenario, with future scalability in mind for all Unified CCE sites.

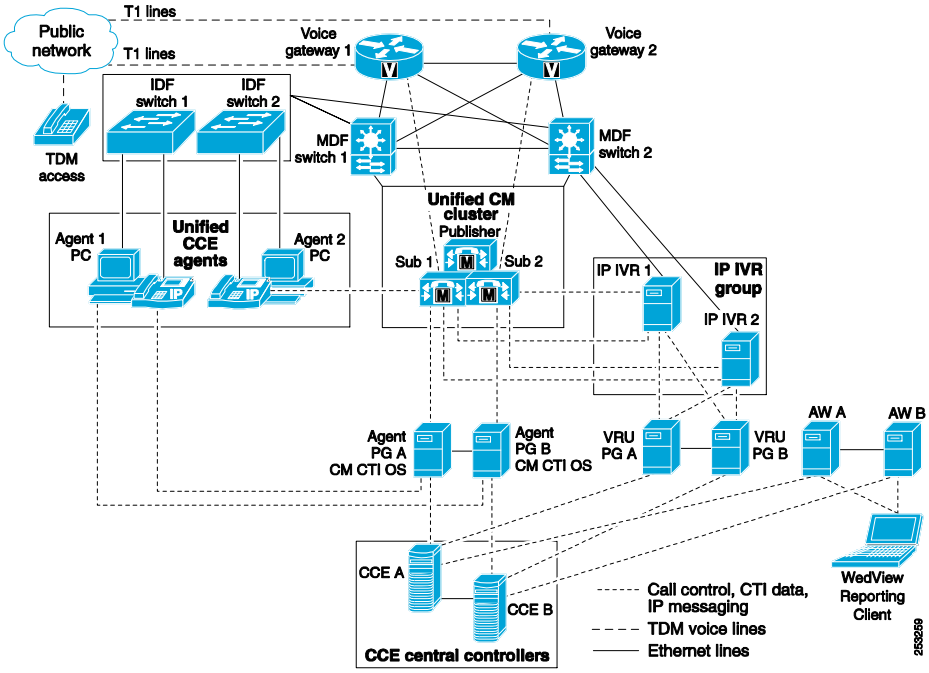
In summary, plan ahead and follow all the design guidelines presented in this guide and in the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

http://www.cisco.com/go/ucsrnd

For assistance in planning and designing your Unified CCE solution, consult your Cisco or certified Partner Systems Engineer (SE).

Figure 3-1 shows a high-level design for a fault-tolerant Unified CCE single-site deployment.

Figure 3-1 Unified CCE Single-Site Design for High Availability



In Figure 3-1, each component in the Unified CCE solution is duplicated with a redundant or duplex component, with the exception of the intermediate distribution frame (IDF) switch for the Unified CCE agents and their phones. The IDF switches do not interconnect with each other, but only with the main distribution frame (MDF) switches, because it is better to distribute the agents among different IDF switches for load balancing and for geographic separation (for example, different building floors or different cities). If an IDF switch fails, route all calls to other available agents in a separate IDF switch or to a Unified IP IVR queue. Follow the design guidelines for a single-site deployment as documented in the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

http://www.cisco.com/go/ucsrnd

If designed correctly for high availability and redundancy, a Unified CCE system can lose half of its core component systems or servers and still be operational. With this type of design, no matter what happens in the Unified CCE system, calls can still be handled in one of the following ways:

- Routed and answered by an available Unified CCE agent using an IP phone or desktop softphone
- Sent to an available Unified IP IVR or Unified CVP port or session
- Answered by the Cisco Unified Communications Manager AutoAttendant or Hunt Group
- Prompted by a Unified IP IVR or Unified CVP announcement that the call center is currently experiencing technical difficulties, and to call back later
- Rerouted to another site with available agents or resources to handle the call

The components in Figure 3-1 can be rearranged to form two connected Unified CCE sites, as illustrated inFigure 3-2.

Figure 3-2 Unified CCE Single-Site Redundancy

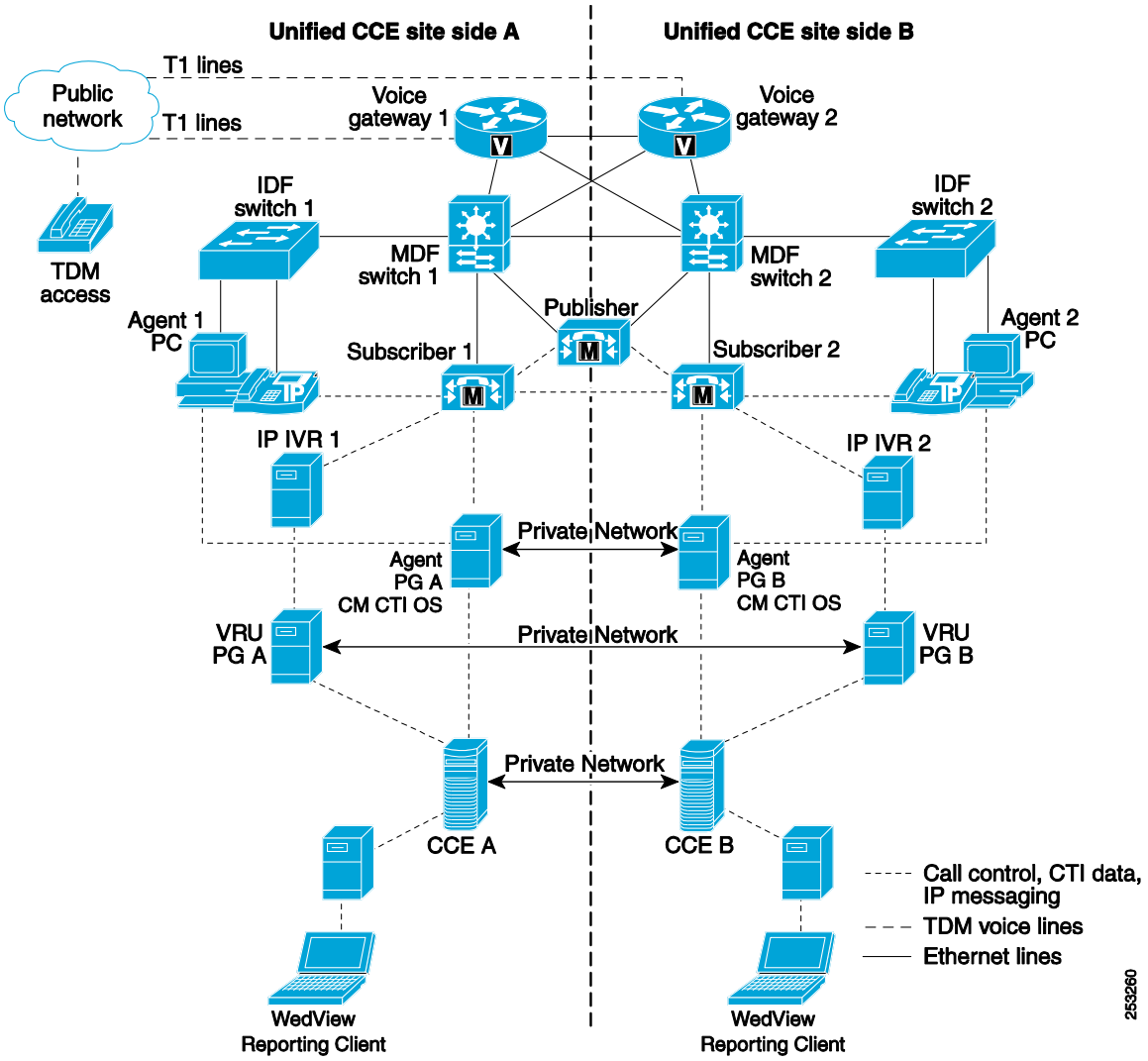


Figure 3-2 emphasizes the redundancy of the single site design in Figure 3-1. Side A and Side B are basically mirror images of each other. In fact, one of the main Unified CCE features to enhance high availability is its ability to add redundant/duplex components that are designed to automatically fail-over and recover without any manual intervention. Core system components with redundant/duplex components are interconnected to provide failure detection of the redundant/duplex system component with the use of TCP keep-alive messages generated every 100 ms over a separate Private Network path. The fault-tolerant design and failure detection/recovery method is described later in this chapter.

Other components in the solution use other types of redundancy strategies. For example, Cisco Unified Communications Manager (Unified CM) uses a cluster design that provides IP phones and devices with multiple Unified CM subscribers (servers) to register when the primary server fails. The devices automatically reconnect to the primary server when it is restored.

The following sections use Figure 3-1 as the model design to discuss issues and features to consider when designing Unified CCE for high availability. These sections use a bottom-up model (from a network model perspective, starting with the physical layer first) that divides the design into segments that can be deployed in separate stages.

Use only duplex (redundant) Unified CM, Unified IP IVR or Unified CVP, and Unified CCE components for all Unified CCE deployments. This chapter assumes that the Unified CCE failover feature is a critical requirement for all deployments, therefore it presents only deployments that use a redundant (duplex) configuration, with each Unified CM cluster having at least one publisher and one subscriber. Additionally, where possible, deploy Unified CCE so that no devices, call processing, or CTI Manager Services are running on the Unified CM publisher.

Data Network Design Considerations

The Unified CCE design shown in Figure 3-3 illustrates the voice call path from the PSTN (public switched telephone network) at the ingress voice gateway to the call reaching a Unified CCE agent. The network infrastructure in the design supports the Unified CCE environment for data and voice traffic. The network, including the PSTN, is the foundation for the Unified CCE solution. If the network is poorly designed to handle failures, then everything in the contact center is prone to failure because all the servers and network devices depend on the network for highly available communications. Therefore, the data and voice networks must be a primary part of your solution design and must be addressed in the early stages for all Unified CCE implementations.

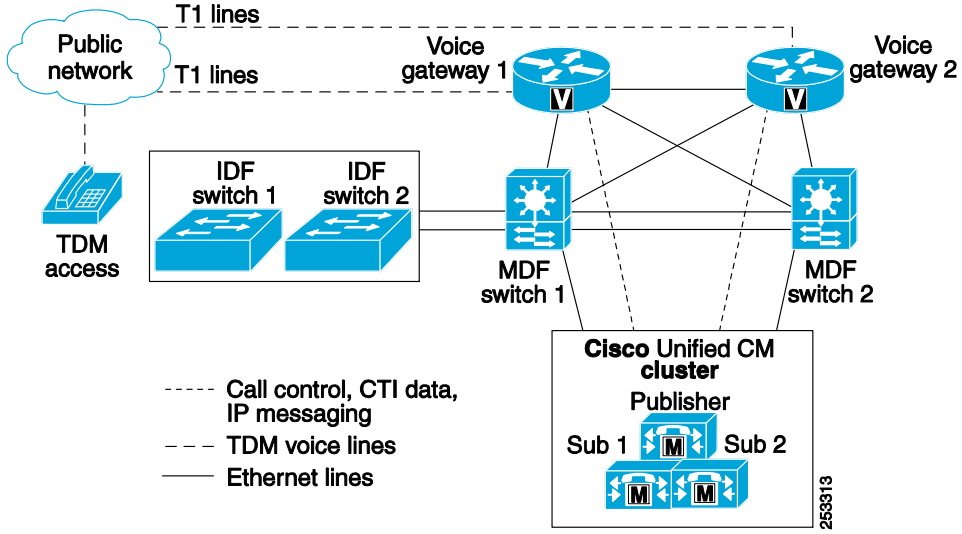
Note Set the NIC card and ethernet switch to 100 MB full duplex for 10/100 links, or set them to auto-negotiate for gigabit links for all the Unified CCE core component servers.

In addition, the choice of voice gateways for a deployment is critical because some protocols offer more call resiliency than others. This chapter provides high-level information on how to configure the voice gateways for high availability with the Unified CCE solution.

For more information on voice gateways and voice networks in general, refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide, available at

http://www.cisco.com/go/ucsrnd

Figure 3-3 High Availability in a Network with Two Voice Gateways and One Unified CM Cluster



Using multiple voice gateways avoids the problem of a single gateway failure causing blockage of all inbound and outgoing calls. In a configuration with two voice gateways and one Unified CM cluster, register each gateway with a different primary Unified CM subscriber to spread the workload across the subscribers in the cluster. Configure each gateway to use another subscriber as a backup in case its primary fails. For details on setting up Unified CM for redundant service and redundancy groups related to call processing, refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide (available at

http://www.cisco.com/go/ucsrnd

With Cisco IOS voice gateways using H.323 or SIP, additional call processing is available by using TCL scripts and additional dial peers if the gateway is unable to reach its Unified CM for call control or call processing instructions. MGCP gateways do not have this built-in functionality, and the trunks that are terminated in these gateways require backup routing or "roll-over service" from the PSTN carrier or service provider to reroute the trunk on failure or no-answer to another gateway or location.

As for sizing the gateway's trunk capacity, it is a good idea to account for failover of the gateways, building in enough excess capacity to handle the maximum busy hour call attempts (BHCA) if one or more voice gateways fail. During the design phase, first decide how many simultaneous voice gateway failures are possible and acceptable for the site. Based upon this requirement, the number of voice gateways used, and the distribution of trunks across those voice gateways, you can determine the total number of trunks required for normal and disaster modes of operation. The more you distribute the trunks over multiple voice gateways, the fewer trunks you will need in a failure mode. However, using more voice gateways or carrier PSTN trunks will increase the cost of the solution, so compare the cost with the benefits of being able to service calls in a gateway failure. The form-factor of the gateway is also a consideration.

As an example, assume a contact center has a maximum BHCA that results in the need for four T1 lines, and the company has a requirement for no call blockage in the event of a single component (voice gateway) failure. If two voice gateways are deployed in this case, then provision each voice gateway with four T1 lines (total of eight). If three voice gateways are deployed, then two T1 lines per voice gateway (total of six) would be enough to achieve the same level of redundancy. If five voice gateways are deployed, then one T1 per voice gateway (total of five) would be enough to achieve the same level of redundancy. Thus, you can reduce the number of T1 lines required by adding more voice gateways and spreading the risk over multiple physical devices.

The operational cost savings of fewer T1 lines might be greater than the one-time capital cost of the additional voice gateways. In addition to the recurring operational costs of the T1 lines, also factor in the carrier charges like the typical one-time installation cost of the T1 lines to ensure that your design accounts for the most cost-effective solution. Every installation has different availability requirements and cost metrics, but using multiple voice gateways is often more cost-effective. Therefore, it is a worthwhile design practice to perform this cost comparison.

After you have determined the number of trunks needed, the PSTN service provider has to configure them so that calls can be terminated onto trunks connected to all of the voice gateways (or at least more than one voice gateway). From the PSTN perspective, if the trunks going to the multiple voice gateways are configured as a single large trunk group, then all calls will automatically be routed to the surviving voice gateways when one voice gateway fails. If all of the trunks are not grouped into a single trunk group within the PSTN, then you must ensure that PSTN rerouting or overflow routing to the other trunk groups is configured for all dialed numbers.

If a voice gateway with a digital interface (T1 or E1) fails, then the PSTN automatically stops sending calls to that voice gateway because the carrier level signaling on the digital circuit has dropped. Loss of carrier level signaling causes the PSTN to busy-out all trunks on that digital circuit, thus preventing the PSTN from routing new calls to the failed voice gateway. When the failed voice gateway comes back on-line and the circuits are back in operation, the PSTN automatically starts delivering calls to that voice gateway again.

With Cisco IOS voice gateways using H.323 or SIP, it is possible for the voice gateway itself to be operational but for its communication paths to the Unified CM servers to be severed (for example, a failed Ethernet connection). If this situation occurs, you can use the **busyout-monitor interface** command to monitor the Ethernet interfaces on a voice gateway. To place a voice port into a busyout monitor state, use the **busyout-monitor interface** *voice-port* configuration command. To remove the busyout-monitor state on the voice port, use the **no** form of this command. As noted previously, these gateways also provide additional processing options if the call control interface is not available from Unified CM to reroute the calls to another site or dialed number or to play a locally stored .wav file to the caller and end the call.

With MGCP-controlled voice gateways, when the voice gateway interface to Unified CM fails, the gateway will look for secondary and tertiary Unified CM subscribers from the redundancy group. The MGCP gateway will automatically fail-over to the other subscribers in the group and periodically check the health of each, marking it as available once it comes back on-line. The gateway will then fail-back to the primary subscriber when all calls are idle or after 24 hours (whichever comes first). If no subscribers are available, the voice gateway automatically busies-out all its trunks. This action prevents new calls from being routed to this voice gateway from the PSTN. When the voice gateway interface to Unified CM homes to the backup subscriber, the trunks are automatically idled and the PSTN begins routing calls to this voice gateway again (assuming the PSTN has not permanently busied-out those trunks). The design practice is to spread the gateways across the Unified CM call processing servers in the cluster to limit the risk of losing all the gateway calls in a call center if the primary subscriber that has all the gateways registered to it fails.

Voice gateways that are used with the Cisco Unified Survivable Remote Site Telephony (SRST) option for Unified CM follow a similar failover process. If the gateway is cut off from the Unified CM that is controlling it, the gateway will fail-over into SRST mode, which drops all voice calls and resets the gateway into SRST mode. Phones re-home to the local SRST gateway for call control, and calls will be processed locally and directed to local phones. While running in SRST mode, it is assumed that the agents also have no CTI connection from their desktops, so they will be seen as not ready within the Unified CCE routing application. Therefore, no calls will be sent to these agents by Unified CCE. When the data connection is re-established to the gateway at the site, the Unified CM will take control of the gateway and phones again, allowing the agents to be reconnected to the Unified CCE.

Unified CM and CTI Manager Design Considerations

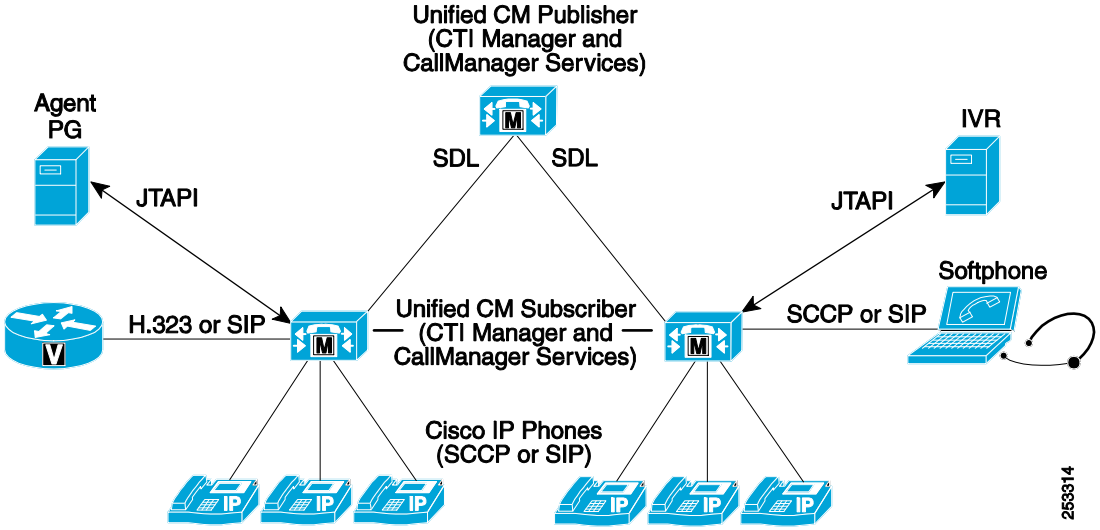
Cisco Unified CM uses CTI Manager, a service that acts as an application broker and abstracts the physical binding of the application to a particular Unified CM server to handle all its CTI resources. (Refer to the *Cisco Unified Communications Solution Reference Network Design (SRND)* guide for further details about the architecture of the CTI Manager.) The CTI Manager and CallManager are two separate services running on a Unified CM server. Some other services running on a Unified CM server include TFTP, Cisco Messaging Interface, and Real-time Information Server (RIS) data collector services.

The main function of the CTI Manager is to accept messages from external CTI applications and send them to the appropriate resource in the Unified CM cluster. The CTI Manager uses the Cisco JTAPI link to communicate with the applications. It acts like a JTAPI messaging router. The JTAPI client library in Cisco Unified CM connects to the CTI Manager instead of connecting directly to the CallManager service. In addition, there can be multiple CTI Manager services running on different Unified CM servers in the cluster that are aware of each other (via the CallManager service, which is explained later in this section). The CTI Manager uses the same Signal Distribution Layer (SDL) signaling mechanism that the Unified CM services in the cluster use to communicate with each other. However, the CTI Manager does not directly communicate with the other CTI Managers in its cluster. (This is also explained later in detail.)

The main function of the CallManager service is to register and monitor all the Cisco Unified Communications devices. It basically acts as a switch for all the Cisco Unified Communications resources and devices in the system, while the CTI Manager service acts as a router for all the CTI application requests for the system devices. Some of the devices that can be controlled by JTAPI that register with the CallManager service include the IP phones, CTI ports, and CTI route points.

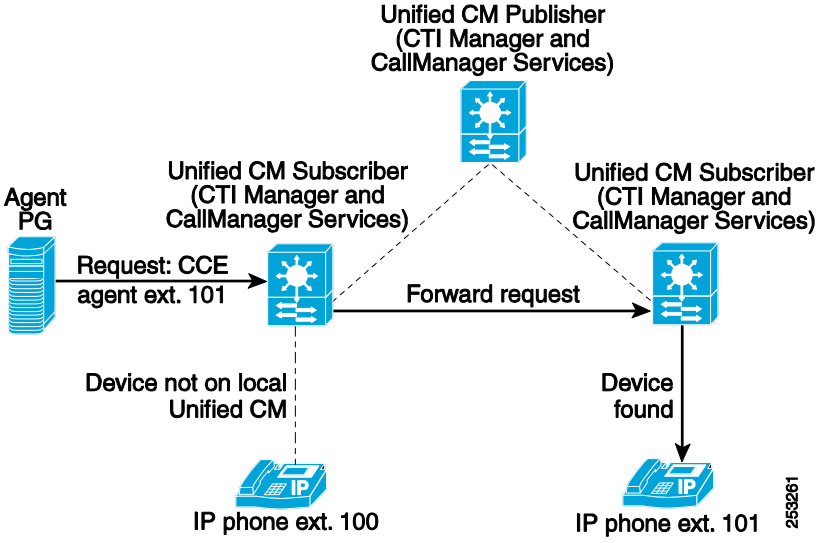
Figure 3-4 illustrates some of the functions of Unified CM and the CTI Manager.

Figure 3-4 Functions of the CallManager and CTI Manager Services



The servers in a Unified CM cluster communicate with each other using the Signal Distribution Layer (SDL) service. SDL signaling is used only by the CallManager service to talk to the other CallManager services to make sure everything is in sync within the Unified CM cluster. The CTI Managers in the cluster are completely independent and do not establish a direct connection with each other. CTI Managers route only the external CTI application requests to the appropriate devices serviced by the local CallManager service on this subscriber. If the device is not resident on its local Unified CM subscriber, then the CallManager service forwards the application request to the appropriate Unified CM in the cluster. Figure 3-5 shows the flow of a device request to another Unified CM in the cluster.

Figure 3-5 CTI Manager Device Request to a Remote Unified CM



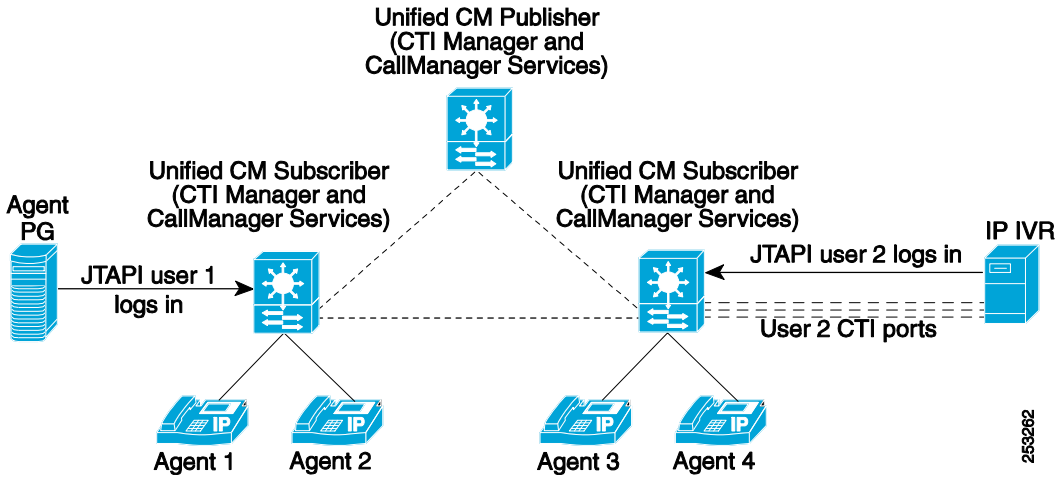
Although it might be tempting to register all of the Unified CCE devices to a single subscriber in the cluster and point the Peripheral Gateway (PG) to that server, this configuration would put a high load on that subscriber. If the PG were to fail in this case, the duplex PG would connect to a different subscriber, and all the CTI Manager messaging would have to be routed across the cluster to the original subscriber. It is important to distribute devices and CTI applications appropriately across all the call processing nodes in the Unified CM cluster to balance the CTI traffic and possible failover conditions.

The external CTI applications use a CTI-enabled user account in Unified CM. They log into the CTI Manager service to establish a connection and assume control of the Unified CM devices associated to this specific CTI-enabled user account, typically referred to as the JTAPI user or PG user. In addition, given that the CTI Managers are independent from each other, any CTI application can connect to any CTI Manager in the cluster to perform its requests. However, because the CTI Managers are independent, one CTI Manager cannot pass the CTI application to another CTI Manager upon failure. If the first CTI Manager fails, the external CTI application must implement the failover mechanism to connect to another CTI Manager in the cluster.

For example, the Agent PG handles failover for the CTI Manager by using its duplex servers, sides A and B, each of which is pointed to a different subscriber in the cluster, and by using the CTI Manager on those subscribers. It is important to note these connections from the PG are managed in hot standby mode, which means only one side of the PG is active at any given time and is connected to the CTI Manager on the subscriber. The PG processes are designed to prevent both sides from trying to be active at the same time to reduce the impact of the CTI application on Unified CM. Additionally, both of the duplex PG servers (Side A and Side B) use the same CTI-enabled JTAPI or PG user to log into the CTI Manager applications. However, only one Unified CM PG side allows the JTAPI user to register and monitor the user devices to conserve system resources in the Unified CM cluster. The other side of the Unified CM PG stays in hot-standby mode, waiting to connect, log in, register, and be activated upon failure of the active side.

Figure 3-6 shows two external CTI applications using the CTI Manager, the Agent PG, and the Unified IP IVR. The Unified CM PG logs into the CTI Manager using the JTAPI account User 1, while the Unified IP IVR uses account User 2. Each external application uses its own specific JTAPI user account and will have different devices registered and monitored by that user. For example, the Unified CM PG (User 1) will monitor all four agent phones and the inbound CTI Route Points, while the Unified IP IVR (User 2) will monitor its CTI Ports and the CTI Route Points used for its JTAPI Triggers. Although multiple applications could monitor the same devices, avoid this method because it can cause race conditions between the applications trying to take control of the same physical device.

Figure 3-6 CTI Application Device Registration



Unified CM CTI applications also add to the device weights on the subscribers, adding memory objects used to monitor registered devices. These monitors are registered on the subscriber that has the connection to the external application. It is a good design practice to distribute these applications to CTI Manager registrations across multiple subscribers to avoid overloading a single subscriber with all of the monitored object tracking.

Perform the design of Unified CM and CTI Manager as the second design stage, right after the network design stage, and perform deployment in this same order. The reason for this order is that the Cisco Unified Communications infrastructure must be in place to dial and receive calls using its devices before you can deploy any telephony applications. Before moving to the next design stage, make sure that a PSTN phone can call an IP phone and that this same IP phone can dial out to a PSTN phone, with all the call survivability capabilities considered for treating these calls. Also keep in mind that the Unified CM cluster design is paramount to the Unified CCE system, and any server failure in a cluster will take down two services (CTI Manager and CallManager), thereby adding an extra load to the remaining servers in the cluster.

Configuring the Unified CCE Peripheral Gateway for CTI Manager Redundancy

To enable Unified CM support for CTI Manager failover in a duplex Unified CCE Peripheral Gateway model, perform the following steps:

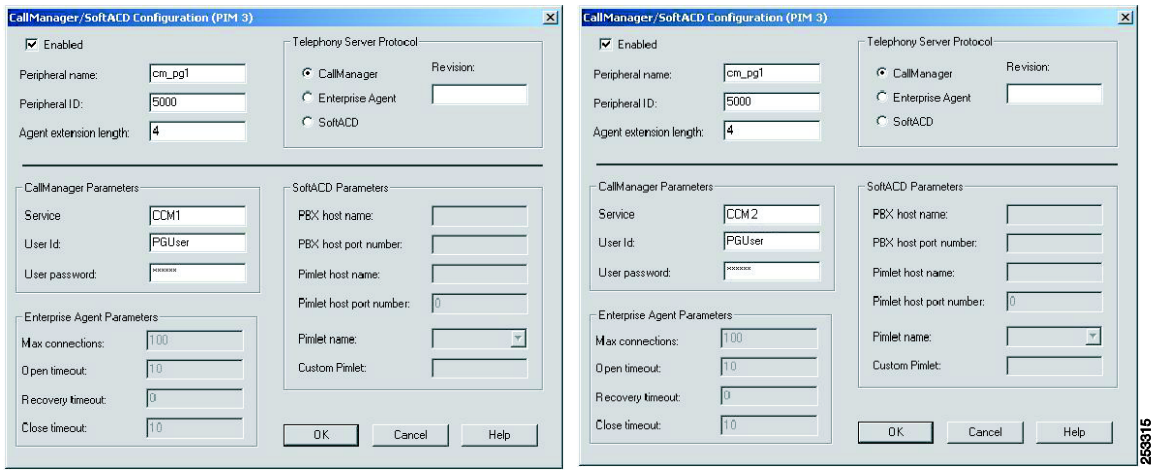
Step 1 Create a Unified CM redundancy group, and add subscribers to the group. (Do not use Publishers and TFTP servers for call processing, device registration, or CTI Manager functions.)

Step 2 Designate two CTI Managers on different subscribers to be used for each side of the duplex Peripheral Gateway (PG), one for PG Side A and one for PG Side B.

Step 3 Assign one of the CTI Managers to be the JTAPI service of the Unified CM PG Side A. (See Figure 3-7.) Note that the setup panel on the left is for Side A of the Peripheral Gateway. It points to the CCM1 subscriber and uses the PGUser CTI-enabled user account on the Unified CM cluster.

Step 4 Assign the second CTI Manager to be the JTAPI service of the Unified CM PG Side B. (See Figure 3-7.) Note that the setup panel on the right is for Side B of the Peripheral Gateway. It points to the CCM2 subscriber and uses the same PGUser CTI-enabled user account on the Unified CM cluster. Both sides of the duplex PG pair must use the same JTAPI user in order to monitor the same devices from either side of the PG pair.

Figure 3-7 Assigning CTI Managers for PG Sides A and B



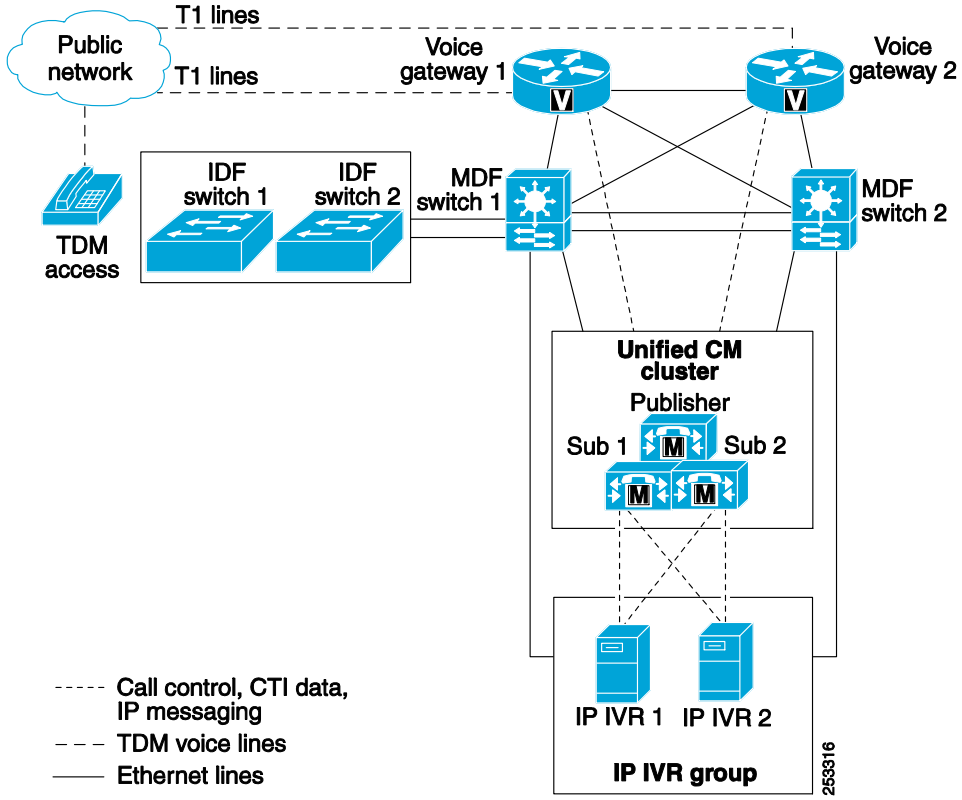
PG side A, Cisco CM PIM 1 PG side B, Cisco CM PIM 1

Unified IP IVR Design Considerations

The JTAPI subsystem in Unified IP IVR can establish connections with two CTI Managers on different subscribers in the Unified CM cluster. This feature enables Unified CCE designs to add Unified IP IVR redundancy at the CTI Manager level, such as the Unified CCE Peripheral Gateway connections. Additionally, deploy multiple, redundant IP-IVR servers in the design and allowing the Unified CCE call routing script to load-balance calls automatically between the available IP-IVR resources.

Figure 3-8 shows two Unified IP IVR servers configured for redundancy within one Unified CM cluster. Configure the Unified IP IVR group so that each server is connected to a different CTI Manager service on different Unified CM subscribers in the cluster for high availability. Using the redundancy feature of the JTAPI subsystem in the Unified IP IVR server, you can implement redundancy by adding the IP addresses or host names of two Unified CMs from the cluster. Then, if one of the Unified CMs fails, the Unified IP IVR associated with that particular Unified CM will fail-over to the second Unified CM.

Figure 3-8 High Availability with Two Unified IP IVR Servers and One Unified CM Cluster



Unified IP IVR High Availability Using Unified CM

You can implement Unified IP IVR port high availability by using any of the following call-forward features in Unified CM:

- Forward Busy — forwards calls to another port or route point when Unified CM detects that the port is busy. This feature can be used to forward calls to another resource when a Unified IP IVR CTI port is busy due to a Unified IP IVR application problem, such as running out of available CTI ports.

- Forward No Answer — forwards calls to another port or route point when Unified CM detects that a port has not picked up a call within the timeout period set in Unified CM. This feature can be used to forward calls to another resource when a Unified IP IVR CTI port is not answering due to a Unified IP IVR application problem.
- Forward on Failure — forwards calls to another port or route point when Unified CM detects a port failure caused by an application error. This feature can be used to forward calls to another resource when a Unified IP IVR CTI port is busy due to a Unified CM application error.

Note When using the call forwarding features to implement high availability of Unified IP IVR ports, avoid creating a loop in the event that all the Unified IP IVR servers are unavailable. Basically, do not establish a path back to the first CTI port that initiated the call forwarding.

Unified IP IVR High Availability Using Unified CCE Call Flow Routing Scripts

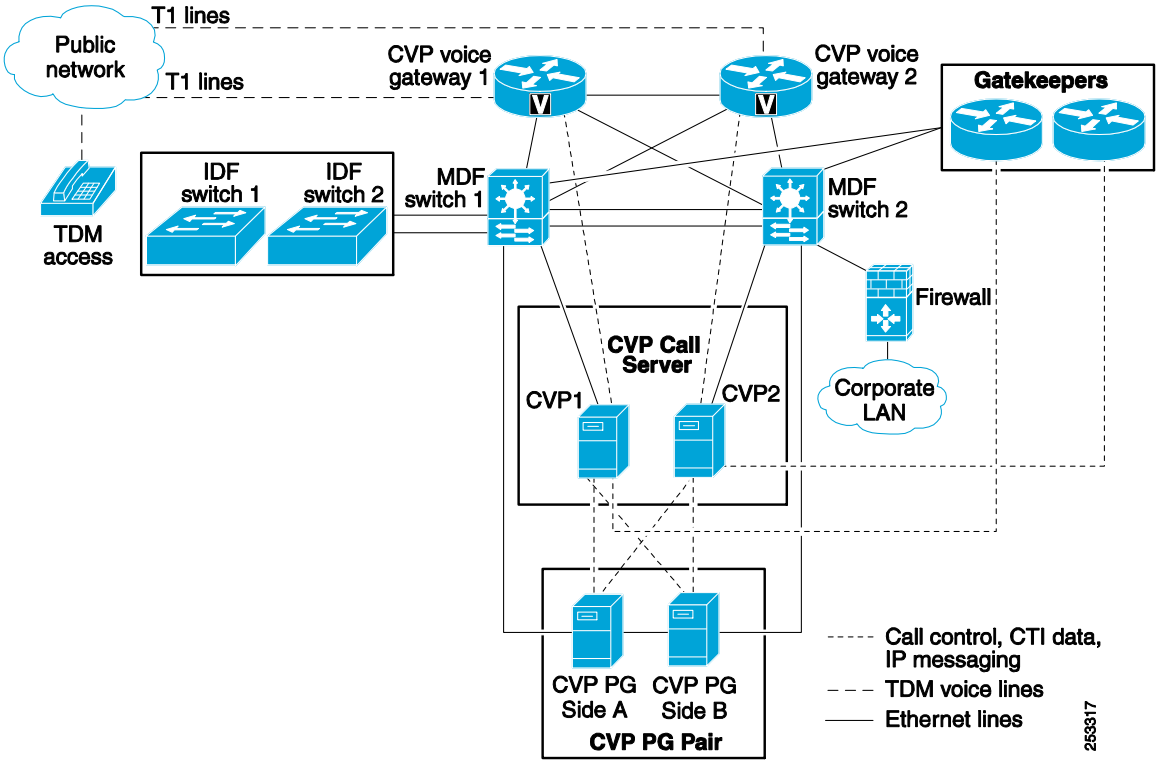
You can implement Unified IP IVR high availability through Unified CCE call flow routing scripts. You can prevent calls from queuing to an inactive Unified IP IVR by using the Unified CCE scripts to check the Unified IP IVR Peripheral Status before sending the calls to it. For example, you can program a Unified CCE script to check if the Unified IP IVR is active by using an IF node or by configuring a Translation Route to the Voice Response Unit (VRU) node (by using the **consider if** field) to select the Unified IP IVR with the most idle ports to distribute the calls evenly on a call-by-call basis. This method can be modified to load-balance ports across multiple Unified IP IVRs, and it can address all of the Unified IP IVRs on the cluster in the same Translation Route or Send to VRU node.

Note All calls at the Unified IP IVR are dropped if the Unified IP IVR server itself fails. It is important to distribute calls across multiple Unified IP IVR servers to minimize the impact of such a failure. In Unified IP IVR), there is a default script to handle cases where the Unified IP IVR loses the link to the IVR Peripheral Gateway, so that the calls are not lost.

Cisco Unified Customer Voice Portal (Unified CVP) Design Considerations

The Unified CVP can be deployed with Unified CCE as an alternative to Unified IP IVR for call treatment and queuing. Unified CVP is different from Unified IP IVR in that it does not rely on Unified CM for JTAPI call control. Unified CVP uses H.323 or SIP for call control and is used in front of Unified CM or other PBX systems as part of a hybrid Unified CCE or migration solution. (See Figure 3-9.)

Figure 3-9 High Availability with Two Unified CVP Call Control Servers Using H.323



Unified CVP uses the following system components:

- Cisco Voice Gateway

  The Cisco Voice Gateway is typically used to terminate TDM PSTN trunks and calls to transform them into IP-based calls on an IP network. Unified CVP uses specific Cisco IOS voice gateways that support H.323 and SIP to enable more flexible call control models outside of the Unified CM MGCP control model. H.323 and SIP protocols enable Unified CVP to integrate with multiple IP and TDM architectures for Unified CCE. Voice gateways controlled by Unified CVP also provide additional functionality using the Cisco IOS built-in Voice Extensible Markup Language (VoiceXML) Browser to provide caller treatment and call queuing on the voice gateway without having to move the call to a physical device such as the IP-IVR or a third-party IVR platform. Unified CVP can also leverage the Media Resource Control Protocol (MRCP) interface of the Cisco IOS voice gateway to add automatic speech recognition (ASR) and text-to-speech (TTS) functions on the gateway as well under Unified CVP control.

- Unified CVP Call Server

  The Unified CVP Call Server provides call control signaling when calls are switched between the ingress gateway and another endpoint gateway or a Unified CCE agent. It also provides the interface to the Unified CCE VRU Peripheral Gateway and translates specific Unified CCE VRU commands into VoiceXML code that is rendered on the Unified CVP Voice Gateway. The Call Server can communicate with the gateways using H.323 or SIP as part of the solution.

- Unified CVP Media Server

The Unified CVP caller treatment is provided either by using ASR/TTS functions via MRCP or with predefined .wav files stored on media servers. The media servers act as web servers and serve up the .wav files to the voice browsers as part of their VoiceXML processing. Media servers can be clustered using the Cisco Content Services Switch (CSS) products, thus allowing multiple media servers to be pooled behind a single URL for access by all the voice browsers in the network.

- Unified CVP VXML Application Server

  Unified CVP provides a VoiceXML service creation environment using an Eclipse toolkit browser, which is hosted in the Unified CVP Call StudioApplication. The Unified CVP VXML server hosts the Unified CVP VoiceXML runtime environment, where the dynamic VoiceXML applications are executed and Java and Web Services calls are processed for external systems and database access.

- H.323 Gatekeepers

  Gatekeepers are used with Unified CVP to register the voice browsers and associate them with specific dialed numbers. When a call comes into the network, the gateway will query the gatekeeper to find out where to send the call based upon the dialed number. The gatekeeper is also aware of the state of the voice browsers and will load-balance calls across them and avoid sending calls to out-of-service voice browsers or ones that have no available sessions.

- SIP Proxy Servers

  SIP Proxy Servers are used with Unified CVP to select voice browsers and associate them with specific dialed numbers. When a call comes into the network, the gateway will query the SIP Proxy Server to find out where to send the call based upon the dialed number.

Availability of Unified CVP can be increased by the following methods:

- Adding redundant Unified CVP Call Servers under control of the Unified CCE Peripheral Gateways, thus allowing the calls to be balanced automatically across multiple Unified CVP Call Servers.
- Adding TCL scripts to the Unified CVP gateway to handle conditions where the gateway cannot contact the Unified CVP Call Server to direct the call correctly.
- Adding gatekeeper redundancy with HSRP or gatekeeper clustering in H.323.
- Adding Cisco Content Server to load-balance .wav file requests across multiple Unified CVP Media Servers and VoiceXML URLs access across multiple servers.

Note Calls in Unified CVP are not dropped if the Unified CVP Call Server or Unified CVP PG fails because they can be redirected to another Unified CVP Call Server on another Unified CVP-controlled gateway as part of the fault-tolerant design using TCL scripts (which are provided with the Unified CVP images) in the voice gateway.

For more information on these options, review the Unified CVP product documentation at

http://www.cisco.com/en/US/products/sw/custcosw/ps1006/tsd_products_support_series_home.html

Cisco Multi-Channel Options with the Cisco Interaction Manager: E-Mail Interaction Manager (EIM) and Web Interaction Manager (WIM)

In 2007, Cisco introduced the replacement for the 5.*x* versions of the Multi-Channel products: Cisco E-Mail Manager (CEM) and Cisco Collaboration Server (CCS). These original products were two separate products that had their own integration methods and web interface for the agents and administrators. The new Cisco Interaction Manager (CIM) platform is a single application that provides

both E-Mail and Web interaction management using a common set of web servers and pages for agents and administrators. The new offering is designed for integration with the Unified CCE platform to provide universal queuing of contacts to agents from different media channels.

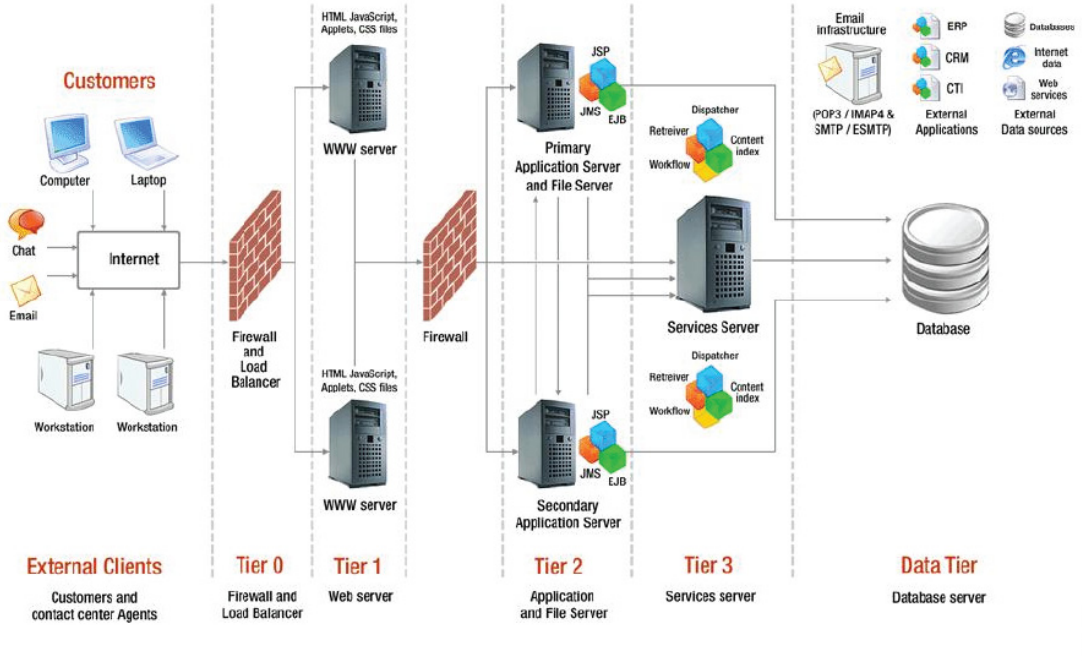
For additional design information about the Interaction Manager platform, refer to the *Cisco Unified Web and E-Mail Interaction Manager Solution Reference Network Design (SRND) Guide for Unified Contact Center Enterprise, Hosted, and ICM*, available at

http://www.cisco.com/en/US/products/ps7236/products_implementation_design_guides_list.html

Cisco Interaction Manager Architecture Overview

The Cisco Interaction Manager has several core components, as illustrated in Figure 3-10.

Figure 3-10 Cisco Interaction Manager Architecture



The architecture is defined by a multi-tiered model, with various components at each of the following levels of the design:

External Clients

Cisco Interaction Manager is a 100% web-based product that agents and end-customers can access using a web browser from their respective desktops.

Agents can access the application using Microsoft Internet Explorer 6.0 or the embedded CAD browser, and customers can access the chat customer console using specific versions of Microsoft IE, Mozilla, Firefox, or Netscape. Cisco Interaction Manager is not supported on agent desktops running in a Citrix terminal services environment.

Tier 0: Firewall and Load Balancer

Agents and customers connect to the application from their respective browsers through a firewall, if so configured for the application.

A load balancer may also be used in case of a distributed installation of the application, so that requests from agents and customers are routed to the least-loaded web servers.

Tier 1: Web Server

The web server is used to serve static content to the browser. Cisco Interaction Manager is designed to be indifferent to the specific type of web server being used, with the single requirement being that the application server vendor must provide a web server plug-in for the corresponding application server.

Tier 2: Application and File Server

The application server is used as a web container (also known as the JSP/Servlet engine) and EJB Container. The core business logic resides in the Business Object Layer, and stored procedures reside on the database server. The business logic residing in JAVA classes is deployed on the application server. The JSP/Servlets interact with the business objects through the business client layer, and these in turn interact with the database to execute some business logic on data present in the database server.

Example: Outbound Task Creation

- A user logs in to the application and creates an outbound task.
- The JSP layer calls the Business Client layer, which interacts with Business Objects residing in the same application server where JSPs/Servlets are deployed.
- The Business Objects execute queries and/or stored procedures residing on the database server.
- Activities are created and stored in database tables.
- The file server is used for storing all email and article attachment files, report templates, and all locale-specific strings used in the application.

Tier 3: Services Server

Cisco Interaction Manager has processes that perform specific business functions, such as fetching emails from a POP server, sending emails to an SMTP server, processing workflows, assigning chats to the agents, and so forth. All services run on the Services server and are managed by the Distributed Service Manager (DSM).

Cisco Interaction Manager facilitates the creation of multiple instances of services with work distributed among the various instances. For example, the service used to retrieve emails could be configured to have multiple instances to retrieve emails from different email addresses. This capability can be used to process increasing volumes of customer interactions coming into a contact center.

Data Tier: Database Server

The data tier includes databases that are SQL-compliant, HTML/XML data-sources, and ultimately Web services that consume and produce SOAP messages. Business objects and data adapters use this layer to extract data from various third-party applications and data sources. This layer also deals with HTML and XML parsing using relevant J2EE-compliant packages to process data in other formats.

Unified CCE Integration

As part of the system integration with Unified CCE, the services server consists of two additional services: the EAAS and the Listener Service, which interact with the Media Routing (MR) PG and Agent PG components of Unified CCE respectively via the Media Routing (MR) and Agent Resource Management (ARM) interfaces.

Additionally, the application server of Cisco Interaction Manager establishes a connection with the Unified CCE Administration & Data server to import relevant configuration data and to map the configuration to Cisco Interaction Manager objects in the Cisco Interaction Manager database. Note that Cisco Interaction Manager does not make use of the Configuration API (ConAPI) interface.

For certain deployments of Unified CCE, the Media Routing (MR) PG of Unified CCE can reside on the services server.

In parent/child configurations, there is no multi-channel routing and integration through the parent Unified ICM. Media Routing PGs need to connect to the child Unified CCE. A separate Cisco Interaction Manager or partition is required for each child.

Likewise, in hosted Unified ICM/CCH environments, there is no multi-channel routing through the Network Application Manager (NAM) layer, and integration is at the individual Customer ICM (CICM) level only. The Media Routing (MR) PGs need to connect to the CICM.

High Availability Considerations for Cisco Interaction Manager

The Cisco Interaction Manager offers high availability options using additional web and application servers and using load balancing equipment to distribute agents and contact work more evenly across the platform as well as to provide for failover in redundancy models.

Load Balancing Considerations

The web service component of a Cisco Interaction Manager deployment can be load balanced to serve a large number of agents accessing the application at the same time. The web (or Web/Application) servers can be configured behind the load balancer with Virtual IP, and an agent can access Cisco Interaction Manager through Virtual IP. Depending on the selected load balancing algorithm, the load balancer will send a request to one of the web/application servers behind it and send a response back to the agent. In this way, from a security perspective, the load balancer serves as a reverse proxy server too.

One of the most essential parameters for configuring a load balancer is to configure it to support sticky sessions with cookie-based persistence. After every scheduled maintenance task, before access is opened for users, verify that all web/application servers are available to share the load. In absence of this, the first web/application server could be overloaded due to the sticky connection feature. With other configurable parameters, you can define a load-balancing algorithm to meet various objectives such as equal load balancing, isolation of the primary web/application server, or sending fewer requests to a low-powered web/application server.

The load balancer monitors the health of all web/application servers in the cluster. If a problem is observed, the load balancer removes the given web/application server from the available pool of servers, thus preventing new web requests from being directed to the problematic web/application server.

Managing Failover

Cisco Interaction Manager supports clustered deployments. This ensures high availability and performance via transparent replication, load balancing, and failover. The following key methods are available for handling failure conditions within a Cisco Interaction Manager and Unified CCE integrated deployment:

- Implementing multiple Web/App servers. If the primary server goes down, the load balancer can help handle the failure through routing requests to alternate Web/App servers. The load balancer detects application server failure and redirects requests to another application server, after which a new user session will be created and users will have to log in again to the Cisco Interaction Manager.

- Allowing servers to be dynamically added or removed from the online cluster to accommodate external changes in demand or internal changes in infrastructure.
- Allowing Cisco Interaction Manager services to fail-over with duplexed Unified CCE components (for example, MR PIM and Agent PIM of the MR PG and Agent PG, respectively) to eliminate downtime of the application in failure circumstances.

The single points of failure in Cisco Interaction Manager include the following.

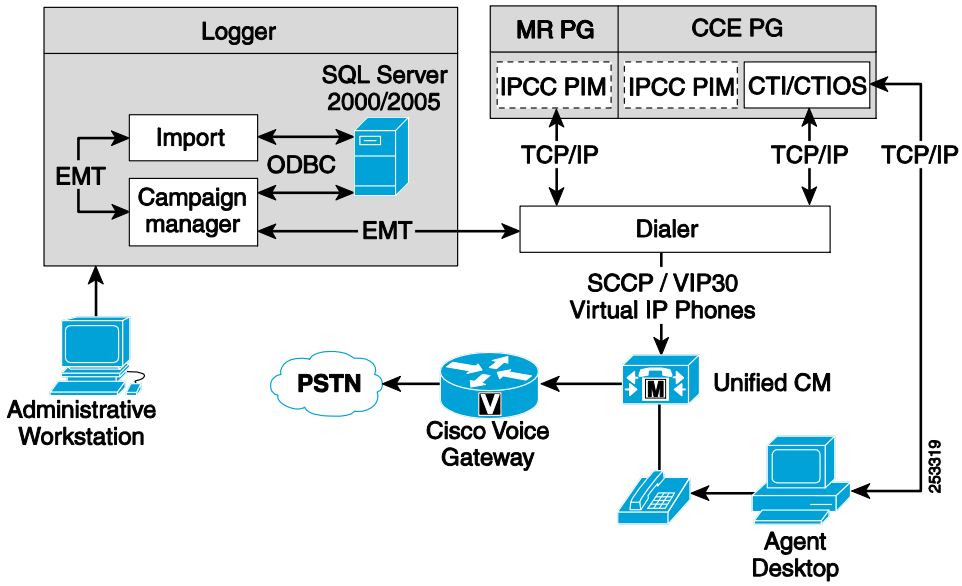
- The JMS server going down
- The Services server going down
- The Database server going down

Cisco Unified Outbound Option Design Considerations

The Cisco Unified Outbound Option provides the ability for Unified CCE to place calls on behalf of agents to customers based upon a predefined campaign. The major components of the Unified Outbound Option are (see Figure 3-11):

- Outbound Option Campaign Manager — A software module that manages the dialing lists and rules associated with the calls to be placed. This software is loaded on the Logger Side A platform and is not redundant; it can be loaded and active on only Logger A of the duplex pair of Loggers in the Unified CCE system.
- Outbound Option Dialer — A software module that performs the dialing tasks on behalf of the Campaign Manager. In Unified CCE, the Outbound Option Dialer emulates a set of IP phones for Unified CM to make the outbound calls, and it detects the called party and manages the interaction tasks with the CTI OS server to transfer the call to an agent. It also interfaces with the Media Routing Peripheral Gateway, and each Dialer has its own peripheral interface manager (PIM) on the Media Routing Peripheral Gateway.
- Media Routing Peripheral Gateway — A software component that is designed to accept route requests from "non-inbound voice" systems such as the Unified Outbound Option or the Multi-Channel products. In the Unified Outbound Option solution, each Dialer communicates with its own peripheral interface manager (PIM) on the Media Routing Peripheral Gateway.

Figure 3-11 Unified CCE Unified Outbound Option



The system can support multiple dialers across the enterprise, all of which are under control of the central Campaign Manager software.

For the new SIP Dialer introduced in Unified CCE Release 8.0, the Dialers operate in a warm standby mode similar to PG fault tolerance model. For more details on this, refer to the Outbound Option chapter.

For the pre-existing SCCP Dialers, although they do not function as a redundant or duplex pair the way a Peripheral Gateway does, with a pair of dialers under control of the Campaign Manager, a failure of one of the dialers can be handled automatically and calls will continue to be placed and processed by the surviving dialer. Any calls that were already connected to agents would remain connected and would experience no impact from the failure.

In all deployments, the Dialers are co-resident on the Unified CCE Peripheral Gateway for Unified CM.

Guidelines for high availability:

- Deploy the Media Routing Peripheral Gateways in duplex pairs.
- Deploy multiple Dialers with one per side of the Duplex Unified CCE Peripheral Gateway, and make use of them in the Campaign Manager to allow for automatic fault recovery to a second Dialer in the event of a failure. For the SCCP Dialer, there are two options with multiple Dialers: a second Dialer can be configured with the same number of ports (100% redundancy), or the ports can be split across the two Dialers since they operate independently and would both be active at the same time. In designs with a small number of Dialer ports, splitting them can impact the performance of the campaign.
- Deploy redundant voice gateways for outbound dialing to ensure that the dialers have enough available trunks to place calls in the event of a voice gateway failure. In some instances where outbound is the primary application, these gateways would be dedicated to outbound calling only.

Peripheral Gateway Design Considerations

The Agent PG uses the Unified CM CTI Manager process to communicate with the Unified CM cluster, with a single Peripheral Interface Manager (PIM) controlling agent phones and CTI route points anywhere in the cluster. The Peripheral Gateway PIM process registers with CTI Manager on one of the Unified CM servers in the cluster, and the CTI Manager accepts all JTAPI requests from the PG for the

cluster. If the phone, route point, or other device being controlled by the PG is not registered to that specific Unified CM server in the cluster, the CTI Manager forwards that request via Unified CM SDL links to the other Unified CM servers in the cluster. There is no need for a PG to connect to multiple Unified CM servers in a cluster.

Multiple PIM Connections to a Single Unified CM Cluster

Although the Agent PG in this document is described as typically having only one PIM process that connects to the Unified CM cluster, the Agent PG can manage multiple PIM interfaces to the same Unified CM cluster, which can be used to create additional peripherals within Unified CCE for two purposes:

- Improving Failover Recovery for Customers with Large Numbers of CTI Route Points
- Scaling the Unified CCE PG Beyond 2,000 Agents per Server

Improving Failover Recovery for Customers with Large Numbers of CTI Route Points

When a Unified CCE PG fails-over, the PIM connection that was previously controlling the Unified CM cluster is disconnected from its CTI Manager, and the duplex or redundant side of the PG will attempt to connect it's PIM to the cluster using a different CTI Manager and Subscriber. This process requires the new PIM connection to register for all of the devices (phones, CTI Route Points, CTI Ports, and so forth) that are controlled by Unified CCE on the cluster. When the PIM makes these registration requests, all of them must be confirmed by Unified CM before the PIM can go into an active state and process calls.

To help recover more quickly, the Unified CCE PG can have a PIM created that is dedicated to the CTI Route Points for the customer, thus allowing this PIM to register for these devices at a rate of approximately five per second and allowing the PIM to activate and respond to calls hitting these CTI Route points faster than if the PIM had to wait for all of the route points, then all the agent phones, and all the CTI ports. This dedicated CTI Route Point PIM could become active several minutes sooner and be able to respond to new inbound calls, directing them to queuing or treatment resources while waiting for the Agent PIM with the phones and CTI Ports to complete the registration process and become active.

This does not provide any additional scaling or other benefits for the design; the only purpose is to allow Unified CM to have the calls on the CTI Route Points serviced faster by this dedicated PIM. Use this only with customers who have more than 250 Route Points, because anything less does not provide a reasonable improvement in recovery time. Additionally,associate only the CTI Route Points that would be serviced by Unified CCE with this PIM, and provide it with its own dedicated CTI-Enabled JTAPI or PGUser specific to the CTI Route Point PIM.

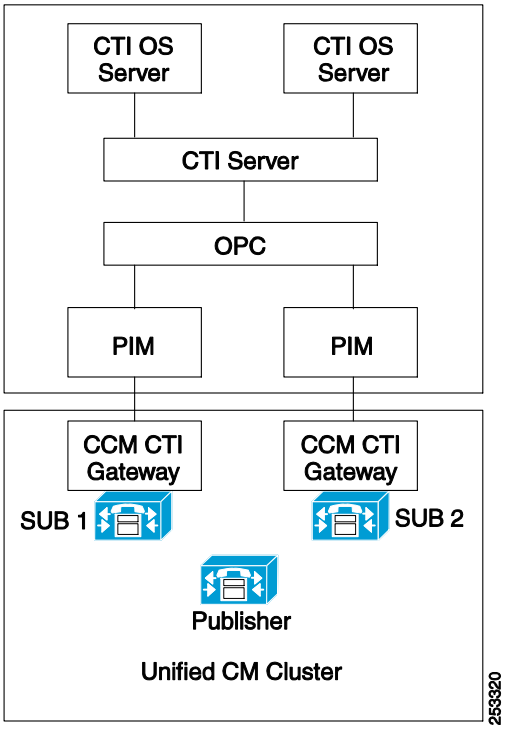
Scaling the Unified CCE PG Beyond 2,000 Agents per Server

In Unified CCE, multiple PIMs in the same physical PG server may be used to connect either to the same Unified CM cluster or to a second Unified CM cluster. This design reduces the physical number of PG servers required in the Unified CCE design. This is different from the recovery strategy for multiple PIMs because both of these PIMs would be configured with up to 2,000 concurrent agents and their related CTI Route Points and CTI Ports as needed to support those agents. The additional PIM will create another Peripheral from the CCE's perspective, which might impact routing and reporting. Additionally, agent teams and supervisors cannot cross peripherals, so careful consideration must be given to which agent groups are allocated to each PIM/Peripheral in such a design.

In designs where Unified CCE is deployed with Unified CVP, the Cisco Unified Communications Sizing Tool might show that the Unified CM cluster can support more than 2,000 total agents; however, the CTI Manager and JTAPI interfaces are tested and supported with a maximum of only 2,000 agents. In order to allow for a design that could have a single Unified CM cluster with more than 2,000 agents, a second Agent PIM can be configured to support the additional agents (up to a total of 4,000 agents per PG).

Figure 3-12 illustrates a single Unified CCE PG with two different PIMs pointing to the same Unified CM cluster.

Figure 3-12 Two PIMs Configured to the Same Unified CM Cluster



Note In order to size the Unified CM cluster properly for Unified CCE, you must use the Cisco Unified Communications Sizing Tool (Unified CST). This tool is available to Cisco partners and employees only, with proper login authentication, at http://tools.cisco.com/cucst.

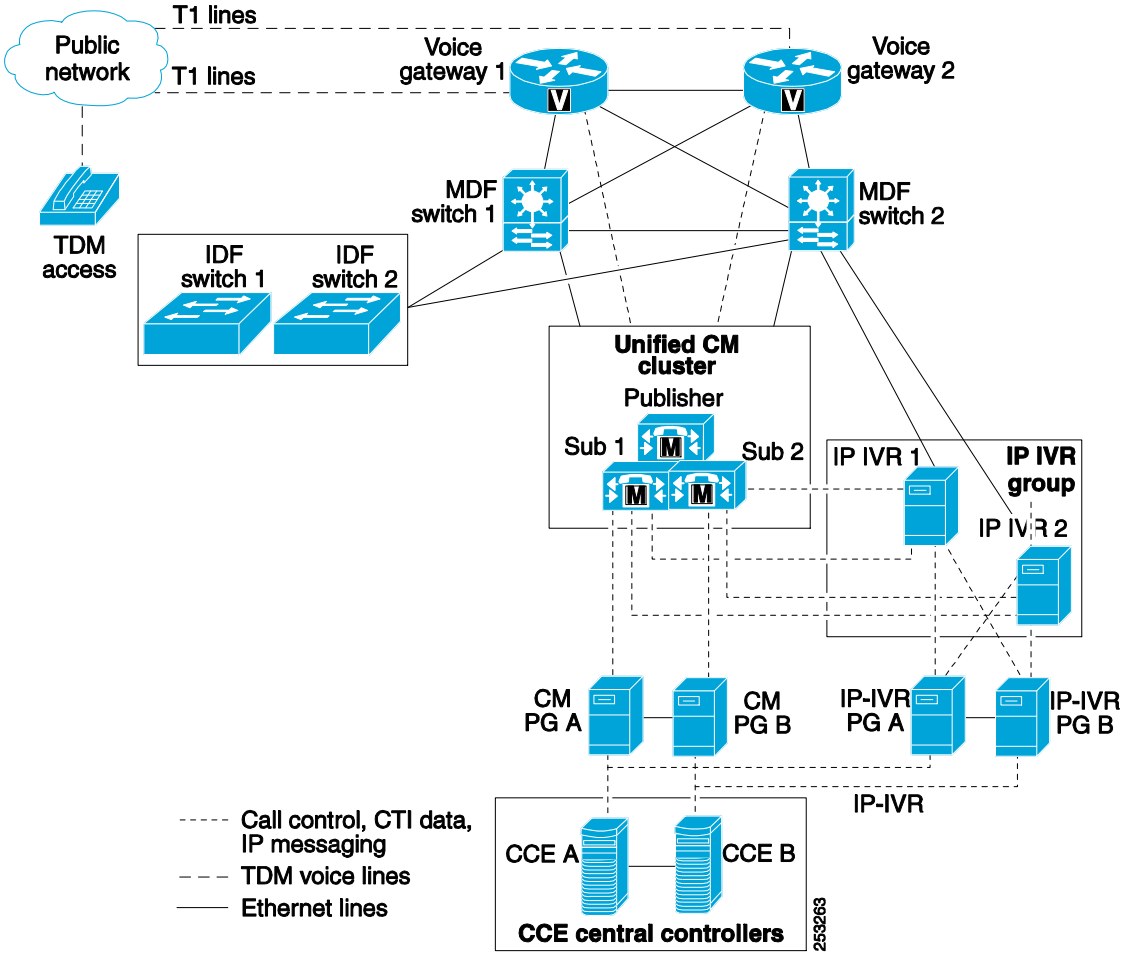
Redundant/Duplex Unified CCE Peripheral Gateway Considerations

Unified CCE Agent PGs are deployed in a redundant/duplex configuration because the PG has only one connection to the Unified CM cluster using a single CTI Manager. If that CTI Manager were to fail, the PG would no long be able to communicate with the Unified CM cluster. Adding a redundant or duplex PG allows Unified CCE to have a second pathway or connection to the Unified CM cluster using a second CTI Manager process on a different Unified CM server in the cluster.

The minimum requirement for Unified CCE high-availability support for CTI Manager and Unified IP IVR is a duplex (redundant) Agent PG environment with one Unified CM cluster containing at least two subscribers. Therefore, the minimum configuration for a Unified CM cluster in this case is one publisher and two subscribers. This minimum configuration ensures that, if the primary subscriber fails, the devices will re-home to the secondary subscriber and not to the publisher for the cluster. (See Figure 3-13.) In smaller systems and labs, Cisco permits a single publisher and single subscriber, which means

if the subscriber fails, then all the devices will be active on the publisher. For specific details about the number of required Unified CM servers, see Sizing Cisco Unified Communications Manager Servers, page 11-1.

Figure 3-13 Unified CCE High Availability with One Unified CM Cluster



To simplify the illustration in Figure 3-13, the Unified CCE Server or Unified CCE Central Controller is represented as a single server, but it is actually a set of servers sized according to the Unified CCE agent count and call volume. The Unified CCE Central Controllers include the following redundant/duplex servers:

- Call Router — The "brain" of the CCE complex that provides intelligent call routing instructions based on real-time conditions it maintains in memory across both the A-Side and B-Side Call Router processes.
- Logger/Database Server — The repository for all configuration and scripting information as well as historical data collected by the system. The Loggers are "paired" with their Call Routers such that Call Router Side A will read and write data only to the Logger A, and the Call Router B will read and write only to the Logger B. Because both sides of the Call Router processes are synchronized, the data written to both Loggers is identical.

In specific deployment models, these two components can be installed on the same physical server, which is referred to as a Rogger, or combined Router/Logger. Refer to the chapter on Sizing Unified CCE Components and Servers, page 10-1, for more details on these specific configurations.

Unified CM JTAPI and Peripheral Gateway Failure Detection

There is a heartbeat mechanism that is used to detect failures between the Unified CM JTAPI link and the Peripheral Gateway. However, unlike the Unified CCE heartbeat methods that use TCP keep-alive messages on the open socket ports, this method uses a specific heartbeat message in the JTAPI messaging protocol between the systems. By default, the heartbeat messages are sent every 30 seconds, and the communications path is reset by the Unified CM or Peripheral Gateway after missing two consecutive heartbeat messages.

This failure detection can be enhanced by using the following procedure to change the heartbeat interval on the JTAPI Gateway client that runs on the Peripheral Gateway:

Step 1 From the Start Menu of the Peripheral Gateway, Select **Programs** -> **Cisco JTAPI** -> **JTAPI Preferences**.

Step 2 Set the **Advanced** -> **Server Heartbeat Interval (sec)** field to 5 seconds.

Do not set this value lower than five seconds because it might impact system performance and trigger an inappropriate failover. This setting determines how often the heartbeats are generated. If it is set to five seconds, the system will fail-over this connection within ten seconds of a loss of network connection because it must detect two consecutive missed heartbeats. The default of 30 seconds takes up to one minute (60 seconds) to take action on a network connection failure.

Because this JTAPI connection between the Peripheral Gateway and Unified CM is supported only locally on the same LAN segment, this is not an issue with latency for this heartbeat value. However, if there are any additional network hops, firewalls, or other devices that cause delay between these two components, then set the heartbeat interval value accordingly to account for this possible condition.

Unified CCE Redundancy Options

Duplex/Redundant Unified CCE servers can be located at the same physical site or can be geographically distributed. This applies specifically to the Central Controller (Call Router/Logger) and Peripheral Gateways.

Under normal operations, the Unified CCE Call Router and Logger/Database Server processes are interconnected through a Private Network connection that is isolated from the Visible/Public Network segment. Configure these servers with a second NIC card for the Private Network connection, and isolate the Private connections from the rest of the Visible/Public Network in their own Cisco Catalyst switch if they are located at the same physical site. If the Central Controllers are geographically separated (located at two different physical sites), under normal operations the same Private Network connections must continue to be isolated and connected between the two physical sites with a separate WAN connection. For normal operations, do not provision this Private Network connection on the same circuits or network gear as the Visible/Public Network WAN connection because that would create a single point of failure that could disable both WAN segments at the same time.

The Unified CCE Peripheral Gateway duplex pair of servers is also interconnected through a Private Network connection that is isolated from the Visible/Public Network segment under normal operations. If the two sides of the duplex pair (Side A and Side B) are both at the same physical site, the Private Network can be created by using an Ethernet Cross-Over Cable between the two servers to interconnect their Private Network NIC cards. If the two servers in the duplex pair are geographically distributed (located at two different physical sites), the Private Network connections must be connected with a separate WAN connection between the two physical sites. Do not provision this Private Network connection on the same circuits or network gear as the Visible/Public Network WAN connection because that would create a single point of failure that could disable both WAN segments at the same time.

For additional details on the Unified ICM network requirements for this connection, refer to the *Installation Guides*, available

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html

For additional details on the Unified CCE network requirements for clustered over the WAN, see the section on IPT: Clustering Over the WAN , page 2-32.

Within the Agent PG, two software processes are run to manage the connectivity to the Unified CM cluster:

- JTAPI Gateway

  The JTAPI Gateway is installed on the PG by downloading it from the Unified CM cluster at the time of the PG installation. This ensures compatibility with the JTAPI and CTI Manager versions in the system. Note that, when either the PG or Unified CM is upgraded, this JTAPI Gateway component must be removed and re-installed on the PG.

  The JTAPI Gateway is started by the PG automatically and runs as a node-managed process, which means that the PG will monitor this process and automatically restart it if it fails for any reason. The JTAPI Gateway handles the low-level JTAPI socket connection protocol and messaging between the PIM and the Unified CM CTI Manager.

- Agent PG Peripheral Interface Manager (PIM)

  The PIM is also a node-managed process and is monitored for unexpected failures and automatically restarted. This process manages the higher-level interface between the Unified CCE and the JTAPI Gateway and Unified CM cluster, requesting specific objects to monitor and handling route requests from the Unified CM cluster.

In a duplex Agent PG environment, both JTAPI services from both Agent PG sides log into the CTI Manager upon initialization. Unified CM PG side A logs into the primary CTI Manager, while PG side B logs into the secondary CTI Manager. However, only the active side of the Unified CM PG registers monitors for phones and CTI route points. The duplex Agent PG pair works in hot-standby mode, with only the active PG side PIM communicating with the Unified CM cluster. The standby side logs into the secondary CTI Manager only to initialize the interface and make it available for a failover. The registration and initialization services of the Unified CM devices take a significant amount of time, and having the CTI Manager available significantly decreases the time for failover.

In duplex PG operation, the side that goes active is the PG side that is first able to connect to the Unified CCE Call Router Server and request configuration information. It is not deterministic based upon the side-A or side-B designation of the PG device, but it depends only upon the ability of the PG to connect to the Call Router, and it ensures that only the PG side that has the best connection to the Call Router will attempt to go active.

The startup process of the PIM requires that all of the CTI route points be registered first, which is done at a rate of 5 route points per second. For systems with a lot of CTI route points (for example, 1000), this process can take as long as 3 minutes to complete before the system will allow any of the agents to log in. This time can be reduced by distributing the devices over multiple PIM interfaces to the Unified CM cluster, as noted above.

In the event that calls arrive at the CTI Route Points in Unified CM but the PIM is not yet fully operational, these calls will fail unless these route points are configured with a recovery number in their "Call Forward on Unregistered" or "Call Forward on Failure" setting. These recovery numbers could be the Cisco Unity voicemail system for the Auto Attendant, or perhaps the company operator position, to ensure that the incoming calls are being answered.

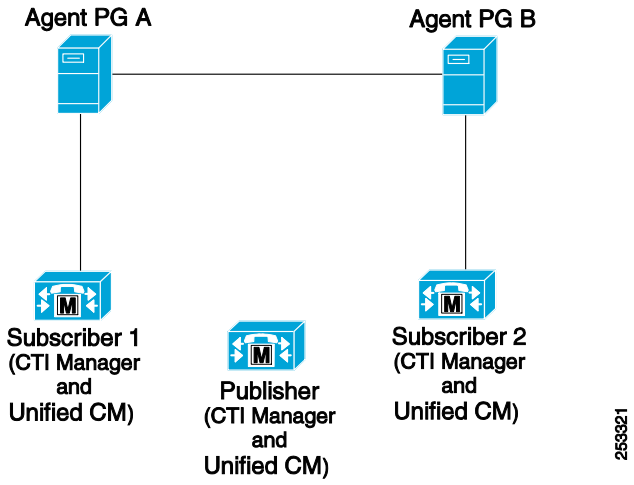
Unified CM Failure Scenarios

A fully redundant Unified CCE system contains no single points of failure. However, there are scenarios where a combination of multiple failures can reduce Unified CCE system functionality and availability. Also, if a component of the Unified CCE solution does not itself support redundancy and failover, existing calls on that component will be dropped. The following failure scenarios have the most impact on high availability, and Unified CM Peripheral Interface Managers (PIMs) cannot activate if either of the following failure scenarios occurs (see Figure 3-14):

- Agent PG/PIM side A and the secondary CTI Manager that services the PG/PIM on side B both fail.
- Agent PG/PIM side B and the primary CTI Manager that services the PG/PIM on side A both fail.

In either of these cases, Unified CCE will not be able to communicate with the Unified CM cluster.

Figure 3-14 Unified CM PGs Cannot Cross-Connect to Backup CTI Managers



Unified CCE Failover Scenarios

This section describes how redundancy works in the following failure scenarios:

- Scenario 1: Unified CM and CTI Manager Fail
- Scenario 2: Agent PG Side A Fails
- Scenario 3: The Unified CM Active Call Processing Subscriber Fails
- Scenario 4: The Unified CM CTI Manager Providing JTAPI Services to the Unified CCE PG Fails

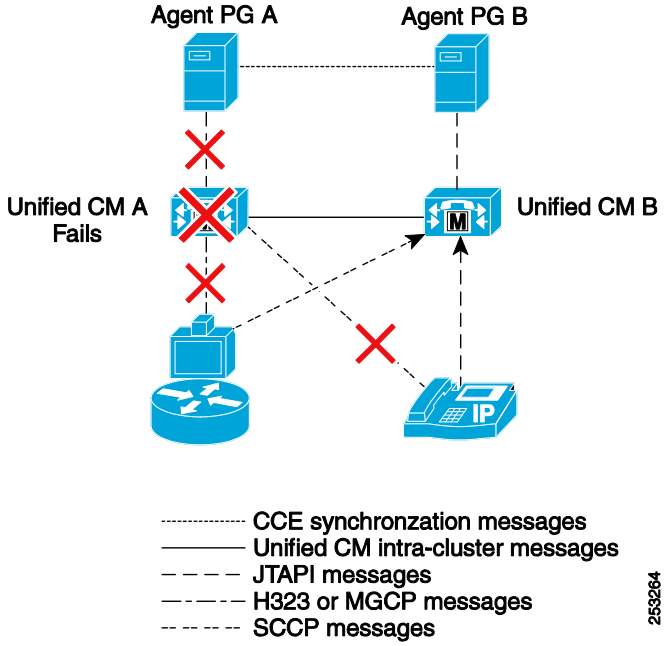
Scenario 1: Unified CM and CTI Manager Fail

Figure 3-15 shows a complete system failure or loss of network connectivity on Cisco Unified CM subscriber A. The CTI Manager and Cisco CallManager services were initially both active on this same server, and Unified CM subscriber A is the primary CTI Manager in this case. The following conditions apply to this scenario:

- All phones and gateways are registered with Unified CM subscriber A as the primary server.
- All phones and gateways are configured to re-home to Unified CM subscriber B (that is, B is the backup server as part of the redundancy group in Unified CM).

- Unified CM subscribers A and B are each running a separate instance of CTI Manager within the same Unified CM cluster.
- When Unified CM subscriber A fails, all registered phones and gateways re-home to Unified CM subscriber B. Calls that are in progress with agent phones will remain active, but the agents will not be able to use phone services such as conference or transfer until they hang up the call and their phone re-registers with the backup subscriber. Although the call stays active, Unified CCE loses visibility to the call and will write a Termination Call Detail (TCD) record to the Unified CCE database for the call at the time of the failure, and no additional call data such as wrap-up codes will be written about the call after that point. Phones that are not active on a call will re-home automatically.
- PG side A detects a failure and induces a failover to PG side B.
- Depending on the configuration of the Peripheral in Unified CCE, the CTI OS or CAD server will keep the agent logged in but will "gray out" their desktop controls until the PG has completed its failover processing. The agents might not have to log in again but might have to manually make themselves "ready" or "available" to ensure they are aware the call processing functionality has been restored.
- PG side B becomes active and registers all dialed numbers and phones, and call processing continues.
- As noted above, when the PG fails-over, the Unified CCE Call Router will write a Termination Call Detail Record (TCD) in the Unified CCE database for any active calls. If the call is still active when the PG fails-over to the other side, a second TCD record will be written for this call as if it were a "new" call in the system and not connected to the prior call that was recorded in the database.
- When Unified CM subscriber A recovers, all idle phones and gateways re-home to it. Active devices wait until they are idle before re-homing to the primary subscriber.
- PG side B remains active, using the CTI Manager on Unified CM subscriber B.
- After recovery from the failure, the PG does *not* fail back to the A side of the duplex pair. All CTI messaging is handled using the CTI Manager on Unified CM subscriber B, which communicates with Unified CM subscriber A to obtain phone state and call information.

Figure 3-15 Scenario 1 - Unified CM and CTI Manager Fail



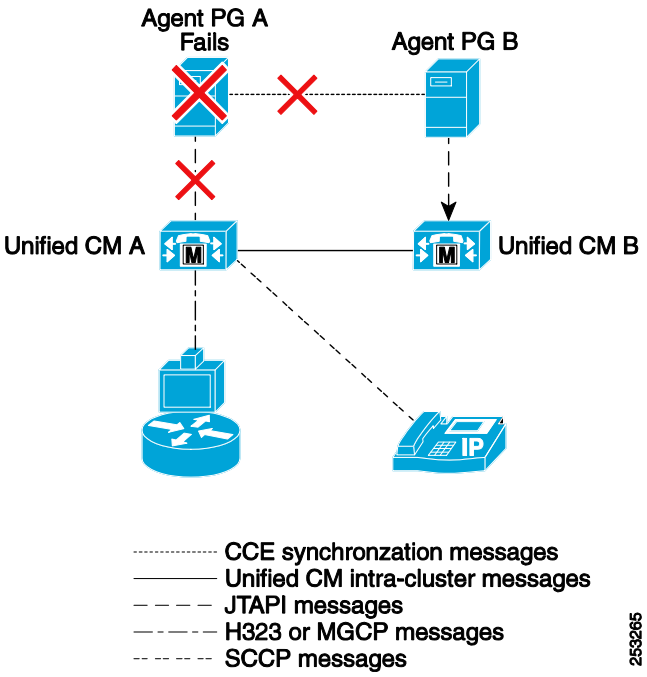
Scenario 2: Agent PG Side A Fails

Figure 3-16 shows a failure on PG side A and a failover to PG side B. All CTI Manager and Unified CM services continue running normally. The following conditions apply to this scenario:

- All phones and gateways are registered with Unified CM subscriber A.
- All phones and gateways are configured to re-home to Unified CM subscriber B (that is, B is the backup server); however, they do not need to re-home as the primary subscriber continues to be functional.
- Unified CM subscribers A and B are each running a local instance of CTI Manager.
- When PG side A fails, PG side B becomes active.
- PG side B registers all dialed numbers and phones, and call processing continues. Phones and gateways stay registered and operational with Unified CM subscriber A; they do not fail-over.
- Agents with calls in progress will stay in progress, but with no third-party call control (conference, transfer, and so forth) available from their agent desktop softphones. Agents that were not on calls may notice their CTI desktop disable their agent state or third-party call control buttons on the desktop during the failover to the B-Side PG. Once the failover is complete, the agent desktop buttons are restored; however the barge in and conference calls will not be rebuilt properly and calls will disappear from the desktop when either of the participants drops out of the call.
- In most cases, after a PG failover, agents whose states are Available or Wrap Up are moved to Available. Alternatively, agents may receive a prompt to log in or to change their state from Not Ready to Available.
- When the PG fails-over, the Unified CCE Call Router will write a Termination Call Detail Record (TCD) in the Unified CCE database for any active calls. If the call is still active when the PG fails-over to the other side, a second TCD record will be written for this call as if it were a "new" call in the system and not connected to the prior call that was recorded in the database.

- When PG side A recovers, PG side B remains active and uses the CTI Manager on Unified CM subscriber B. The PG will not fail-back to the A-Side, and call processing will continue on the PG Side B.

Figure 3-16 Scenario 2 - Agent PG Side A Fails



Scenario 3: The Unified CM Active Call Processing Subscriber Fails

Figure 3-17 shows a failure on Unified CM active call processing subscriber A. In this model, the subscriber is actively processing calls and controlling devices but does not provide the CTI Manager connection to the Unified CCE PG. The CTI Manager services are running on all the Unified CM subscribers in the cluster, but only subscribers C and D are configured to communicate with the Unified CCE Peripheral Gateway.

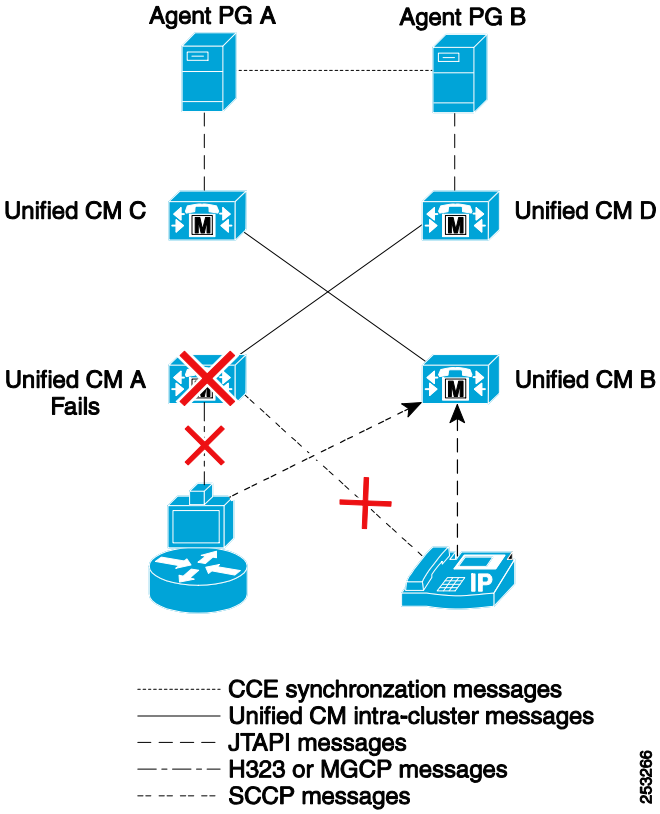
The following conditions apply to this scenario:

- All phones and gateways are registered with Unified CM subscriber A.
- All phones and gateways are configured to re-home to Unified CM subscriber B (that is, B is the backup server).
- Unified CM subscribers C and D are each running a local instance of CTI Manager to provide JTAPI services for the Unified CCE PGs.
- If Unified CM subscriber A fails, phones and gateways re-home to the backup Unified CM subscriber B.
- PG side A remains connected and active, with a CTI Manager connection on Unified CM subscriber C. It does not fail-over because the JTAPI-to-CTI Manager connection has not failed. However, it will see the phones and devices being unregistered from Unified CM subscriber A (where they were registered) and will then be notified of these devices being re-registered on Unified CM subscriber B automatically. During the time that the agent phones are not registered, the PG will disable the

agent CTI desktops to prevent the agents from attempting to use the system while their phones are not actively registered with a Unified CM subscriber. Also, they will be "logged out" by the system during this transition to avoid routing calls to them as well.

- Call processing continues for any devices not registered to Unified CM subscriber A. Call processing also continues for those devices on subscriber A when they are re-registered with their backup subscriber.
- Calls in progress on phones registered to Unified CM subscriber A will continue; however, the agent desktop will be disabled to prevent any conference, transfer, or other third-party call control during the failover. After the agent disconnects the active call, that agent's phone will re-register with the backup subscriber.
- As noted above, when the Unified CM subscriber A fails, the calls in progress stay active; however, Unified CCE loses control and track of those calls because the phone has not re-homed (re-registered) with the backup subscriber in the cluster. In fact, the phone will not re-home until after the current call is completed. The Unified CCE Call Router will write a Termination Call Detail Record (TCD) in the Unified CCE database for calls that were active at the time of the subscriber failure, with call statistics up to the time of the failure and loss of control. Any additional call information (statistics, call wrap-up data, and so forth) will not be written to the Unified CCE database.
- When Unified CM subscriber A recovers, phones and gateways re-home to it. This re-homing can be set up on Unified CM to gracefully return groups of phones and devices over time or to require manual intervention during a maintenance window to minimize the impact to the call center. During this re-homing process, the CTI Manager service will notify the Unified CCE Peripheral Gateway of the phones being unregistered from the backup Unified CM subscriber B and re-registered with the original Unified CM subscriber A.
- Call processing continues normally after the phones and devices have returned to their original subscriber.

Figure 3-17 Scenario 3 - Only the Primary Unified CM Subscriber Fails



Scenario 4: The Unified CM CTI Manager Providing JTAPI Services to the Unified CCE PG Fails

Figure 3-18 shows a CTI Manager service failure on Unified CM subscriber C that is used to communicate with the Unified CCE PG. The CTI Manager services are running on all the Unified CM subscribers in the cluster, but only subscribers C and D are configured to connect to the Unified CCE PGs. During this failure, the PG will detect the loss of the JTAPI connection and fail-over to the redundant/duplex PG side.

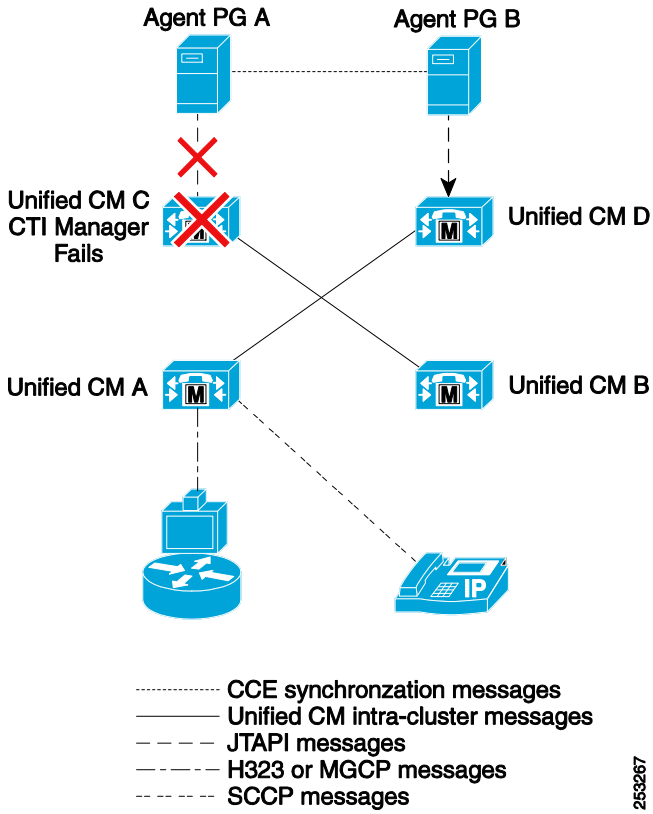
The following conditions apply to this scenario:

- All phones and gateways are registered with Unified CM subscriber A.
- All phones and gateways are configured to re-home to Unified CM subscriber B (that is, B is the backup server). In this case they will not re-home because subscriber A is still functional.
- Unified CM subscribers C and D are each running a local instance of CTI Manager and are designed to connect to the Unified CCE PGs.
- If the Unified CM CTI Manager service on subscriber C fails, the PG side A detects a failure of the CTI Manager service and induces a failover to PG side B.
- PG side B registers all dialed numbers and phones with the Unified CM CTI Manager service on subscriber D, and call processing continues.
- Agents with calls in progress will stay in progress, but with no third-party call control (conference, transfer, and so forth) available from their agent desktop softphones. After an agent disconnects from all calls, that agent's desktop functionality is restored. Although the call stays active, Unified

CCE loses visibility to the call and will write a Termination Call Detail (TCD) record to the Unified CCE database for the call at the time of the failure, and no additional call data such as wrap-up codes will be written about the call after that point.

- When the Unified CM CTI Manager service on subscriber C recovers, PG side B continues to be active and uses the CTI Manager service on Unified CM subscriber D. The PG does not fail-back in this model.

Figure 3-18 Scenario 4 - Only the Unified CM CTI Manager Service Fails



Unified CCE Scenarios for Clustering over the WAN

Unified CCE can also be overlaid with the Unified CM design model for clustering over the WAN, which allows for high availability of Unified CM resources across multiple locations and data center locations. There are a number of specific design requirements for Unified CM to support this deployment model, and Unified CCE adds its own specific requirements and new failover considerations to the model.

Specific testing has been performed to identify the design requirements and failover scenarios. The success of this design model relies on specific network configuration and setup, and the network must be monitored and maintained. The component failure scenarios noted previously (see Unified CCE Failover Scenarios) are still valid in this model, and the additional failure scenarios for this model include:

- Scenario 1: Unified CCE Central Controller or Peripheral Gateway Private Network Failure
- Scenario 2: Visible Network Failure
- Scenario 3: Visible and Private Networks Both Fail (Dual Failure)
- Scenario 4: Unified CCE Agent Site WAN (Visible Network) Failure

Note The terms *public network* and *visible network* are used interchangeably throughout this document.

Scenario 1: Unified CCE Central Controller or Peripheral Gateway Private Network Failure

In clustering over the WAN with Unified CCE, provide a separate private network connection between the geographically distributed Central Controller (Call Router/Logger) and the split Peripheral Gateway pair to maintain state and synchronization between the sides of the system.

To understand this scenario fully, a brief review of the Unified CCE Fault Tolerant architecture is warranted. On each call router, there is a process known as the Message Delivery Service (MDS), which delivers messages to and from local processes such as router.exe and which handles synchronization of messages to *both* call routers. For example, if a route request comes from the carrier or any routing client to side A, MDS ensures that both call routers receive the request. MDS also handles the duplicate output messages.

The MDS process ensures that duplex Unified CCE sides are functioning in a synchronized execution, fault tolerance method. Both routers are executing everything in lockstep, based on input the router receives from MDS. Because of this synchronized execution method, the MDS processes must always be in communication with each other over the private network. They use TCP keep-alive messages generated every 100 ms to ensure the health of the redundant mate or the other side. Missing five consecutive TCP keep-alive messages indicates to Unified CCE that the link or the remote partner system might have failed.

When running duplexed Unified CCE sides for all production system, one MDS will be the enabled synchronizer and will be in a *paired-enabled* state. Its partner will be the disabled synchronizer and is said to be *paired-disabled*. Whenever the sides are running synchronized, the side A MDS will be the enabled synchronizer in paired-enabled state. Its partner, side B, will be the disabled synchronizer in paired-disabled state. The enabled synchronizer sets the ordering of input messages to the router and also maintains the master clock for the Unified CCE system.

If the private network fails between the Unified CCE Central Controllers, the following conditions apply:

- The Call Routers detects the failure by missing five consecutive TCP keep-alive messages. The currently enabled side (side A in most cases) transitions to an isolated-enabled state and continues to function as long as it is in communication with at least half of the PGs configured in the system.
- The paired-disabled side (side B in most cases) transitions to an isolated-disabled state. This side will then check for device majority. If it is not communicating with either an Active or Idle DMP to more than half of the configured PGs in the system, it will stop processing and stay disabled.
- If the B-Side has device majority (an Active or Idle connection to more than half the configured PGs), it will transition to a "Testing" state and send "Test Other Side" (TOS) messages to each PG. This message is used to ask the PG if it can see the Call Router on the other side (in this case, Router A).
- As soon as any (even one) PG responds to the TOS message that the A-Side is still enabled, Router B remains in the Isolated-Disabled state and goes idle. Logger B will also go idle, as will all the DMP connections to the PGs for Router B. All call processing will continue on Side A without impact.
- If all of the PGs reply that Side A is down, or not reachable, the B-Side Call Router would re-initialize in simplex mode (isolated-enabled) and take over all routing for Unified CCE.

- There is no impact to the agents, calls in progress, or calls in queue. The system can continue to function normally; however; the Call Routers will be in simplex mode until the private network link is restored.

Additional Considerations

The Call Routers are "paired" with the Loggers and can read/write only to their own Logger for configuration and historical data over the Private Network locally. In the event that the failure is caused by the loss of a Private NIC card in the Call Router, and that Call Router is the enabled side, it will not be able to write any historical data to the Logger nor will any configuration changes be able to be made to the Logger database.

The Private NIC in the Call Router is also used in some cases to communicate with carrier-based Pre-Routing Network or SS7 interfaces. If the Private NIC fails, there would be no way to access these services either.

If there is an even number of PGs checked off in the Call Router Setup, and only half of the PGs are available, then only Side A will run. For the B-Side to be operational during a private network failure, it must be able to communicate with more than half of the PGs in the system.

It is important to maintain the configuration so that "extra" PGs or PGs that are no longer on the network are removed from the Call Router Setup panels to avoid problems with determination of device majority for PGs that no longer exist.

If the private network fails between the Unified CM Peripheral Gateways, the following conditions apply:

- The Peripheral Gateway sides detect a failure if they miss five consecutive TCP keep-alive messages, and they follow a process similar to the Call Routers, leveraging the MDS process when handling a private link failure. As with the Central Controllers, one MDS process is the enabled synchronizer and its redundant side is the disabled synchronizer. When running redundant PGs, the A side will always be the enabled synchronizer.
- After detecting the failure, the disabled synchronizer (side B) initiates a test of its peer synchronizer via the TOS procedure on the Public or Visible Network connection. If PG side B receives a TOS response stating that the A side synchronizer is enabled or active, then the B side immediately goes out of service, leaving the A side to run in simplex mode until the Private Network connection is restored. The PIM, OPC, and CTI SVR processes become active on PG side A, if not already in that state, and the CTI OS Server process still remains active on both sides as long as the PG side B server is healthy. If the B side does not receive a message stating that the A side is enabled, then side B continues to run in simplex mode and the PIM, OPC, and CTI SVR processes become active on PG side B if not already in that state. This condition occurs only if the PG side A server is truly down or unreachable due to a double failure of visible and private network paths.
- There is no impact to the agents, calls in progress, or calls in queue because the agents stay connected to their already established CTI OS Server process connection. The system can continue to function normally; however, the PGs will be in simplex mode until the private network link is restored.

If the two private network connections are combined into one link, the failures follow the same path; however, the system runs in simplex mode on both the Call Router and the Peripheral Gateway. If a second failure were to occur at that point, the system could lose some or all of the call routing and ACD functionality.

Scenario 2: Visible Network Failure

The visible network in this design model is the network path between the data center locations where the main system components (Unified CM subscribers, Peripheral Gateways, Unified IP IVR/Unified CVP components, and so forth) are located. This network is used to carry all the voice traffic (RTP stream and call control signaling), Unified CCE CTI (call control signaling) traffic, as well as all typical data network traffic between the sites. In order to meet the requirements of Unified CM clustering over the WAN, this link must be highly available with very low latency and sufficient bandwidth. This link is critical to the Unified CCE design because it is part of the fault-tolerant design of the system, and it must be highly resilient as well:

- The highly available (HA) WAN between the central sites must be fully redundant with no single point of failure. (For information regarding site-to-site redundancy options, refer to the WAN infrastructure and QoS design guides available at http://www.cisco.com/go/designzone.) In case of partial failure of the highly available WAN, the redundant link must be capable of handling the full central-site load with all QoS parameters. For more information, see the section onBandwidth Requirements for Unified CCE Clustering Over the WAN , page 12-20.
- A highly available (HA) WAN using point-to-point technology is best implemented across two separate carriers, but this is not necessary when using a ring technology.

If the visible network fails between the data center locations, the following conditions apply:

- The Unified CM subscribers will detect the failure and continue to function locally, with no impact to local call processing and call control. However, any calls that were set up over this WAN link will fail with the link.
- The Unified CCE Call Routers will detect the failure because the normal flow of TCP keep-alives from the remote Peripheral Gateways will stop. Likewise, the Peripheral Gateways will detect this failure by the loss of TCP keep-alives from the remote Call Routers. The Peripheral Gateways will automatically realign their data communications to the local Call Router, and the local Call Router will then use the private network to pass data to the Call Router on the other side to continue call processing. This does not cause a failover of the Peripheral Gateway or the Call Router.
- Half the agents or more might be affected by this failure under the following circumstances:
  - If the agent desktop (Cisco Agent Desktop or CTI OS) is registered to the Peripheral Gateway on side A of the system but the physical phone is registered to side B of the Unified CM cluster.

    Under normal circumstances, the phone events would be passed from side B to side A over the visible network via the CTI Manager Service to present these events to the side A Peripheral Gateway. The visible network failure will not force the IP phone to re-home to side A of the cluster, and the phone will remain operational on the isolated side B. The Peripheral Gateway will no longer be able to see this phone, and the agent will be logged out of Unified CCE automatically because the system can no longer direct calls to the agent's phone.
  - If the agent desktop (Cisco Agent Desktop or CTI OS) and IP phone are both registered to side A of the Peripheral Gateway and Unified CM, but the phone is reset and it re-registers to a side B of the Unified CM subscriber.

    If the IP phone re-homes or is manually reset and forced to register to side B of a Unified CM subscriber, the Unified CM subscriber on side A that is providing the CTI Manager service to the local Peripheral Gateway will unregister the phone and remove it from service. Because the visible network is down, the remote Unified CM subscriber at side B cannot send the phone registration event to the remote Peripheral Gateway. Unified CCE will log out this agent because it can no longer control the phone for the agent.
  - If the agent desktop (CTI Toolkit Agent Desktop or Cisco Agent Desktop) is registered to the CTI OS Server at the side-B site but the active Peripheral Gateway side is at the side-A site.

Under normal operation, the CTI Toolkit Agent Desktop will load-balance their connections to the CTI OS Server pair. At any given time, half the agent connections would be on a CTI OS server that has to cross the visible network to connect to the active Peripheral Gateway CTI Server (CG). When the visible network fails, the CTI OS Server detects the loss of connection with the remote Peripheral Gateway CTI Server (CG) and disconnects the active agent desktop clients to force them to re-home to the redundant CTI OS Server at the remote site. The CTI Toolkit Agent Desktop is aware of the redundant CTI OS server and will automatically use this server. During this transition, the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to the redundant CTI OS server. (The agent may be logged out or put into not-read state, depending upon the /LOAD parameter defined for the Unified CM Peripheral Gateway in Unified CCE Config Manager).

Scenario 3: Visible and Private Networks Both Fail (Dual Failure)

Individually, the private and visible networks can fail with limited impact to the Unified CCE agents and calls. However, if both of these networks fail at the same time, the system will be reduced to very limited functionality. This failure is considered catastrophic and can be avoided by careful WAN design, with backup and resiliency built into the design.

If both the visible and private networks fail at the same time, the following conditions apply:

- The Unified CM subscribers will detect the failure and continue to function locally, with no impact to local call processing and call control. However, any calls that were set up and are sending the active voice path media over the visible WAN link will fail with the link. When the call fails, the Unified CCE PG will see the call drop and will write a Termination Call Detail (TCD) record in the Unified CCE database for that call at the time it is dropped.
- The Call Routers and Peripheral Gateways will detect the private network failure after missing five consecutive TCP keep-alive messages. These TCP keep-alive messages are generated every 100 ms, and the failure will be detected within about 500 ms on this link.
- The Call Routers will attempt to contact their Peripheral Gateways with the test-other-side message to determine if the failure was a network issue or if the remote Call Router had failed and was no longer able to send TCP keep-alive messages. The Call Routers determine which side will continue to be active (typically, this would be the A-Side of the system because it is the side with the most active Peripheral Gateway connections), and that side will stay active in simplex mode while the remote Call Router and PGs will be in isolated-disabled mode. The Call Routers will send a message to the Peripheral Gateways to realign their data feeds to the active Call Router only.
- The Peripheral Gateways will determine which side has the active Unified CM connection. However, it will also consider the state of the Call Router, and the Peripheral Gateway will not remain active if it is not able to connect to an active Call Router. Typically, this will force the A-Side PGs into active simplex enabled mode and the B-Side into isolated-disabled mode.
- The surviving Call Router and Peripheral Gateways will detect the failure of the visible network by the loss of TCP keep-alives on the visible network. These keep-alives are sent every 400 ms, so it can take up to two seconds before this failure is detected.
- The Call Router will be able to see only the local Peripheral Gateways, which are those used to control local Unified IP IVRs or Unified CVP Call Servers and the local half of the Unified CM cluster. The remote Unified IP IVRs or Unified CVP Call Servers will be off-line with no Unified CCE Call Control via the GED-125 IVR PG interface. The Unified CCE Call Routing Scripts automatically routes around these off-line devices using the peripheral-on-line status checks. Calls that were in progress in the off-line IP-IVRs will either drop or use the local default script in the IP-IVR or the Call Forward on Error settings in Unified CM. Calls under Unified CVP control from the off-line Call Servers will get treatment from the survivability TCL script in their ingress voice

gateways. For calls that were in progress but are no longer visible to Unified CCE, a Termination Call Detail (TCD) record is written to the Unified CCE database for the call data up to the time of the failure. If the default or survivability scripts redirect the calls to another active Unified CCE component, the call will appear as a "new call" to the system, with no relationship to the original call for reporting or tracking purposes.

- Any new calls that come into the disabled side will not be routed by Unified CCE, but they can be redirected or handled using standard Unified CM redirect on failure for their CTI route points or the Unified CVP survivability TCL script in the ingress voice gateways.
- Agents will be impacted as noted above if their IP phones are registered to the side of the Unified CM cluster opposite the location of their active Peripheral Gateway and CTI OS Server connection. Only agents that were active on the surviving side of the Peripheral Gateway with phones registered locally to that site will not be impacted.

At this point, the Call Router and Unified CM Peripheral Gateway will run in simplex mode, and the system will accept new calls from only the surviving side for Unified CCE call treatment. The Unified IP IVR/Unified CVP functionality will also be limited to the surviving side as well.

Scenario 4: Unified CCE Agent Site WAN (Visible Network) Failure

The Unified CCE design model for clustering over the WAN assumes the Unified CCE agents are remotely located at multiple sites connected by the visible WAN. Each agent location requires WAN connectivity to both of the data center locations across the visible WAN where the Unified CM and Unified CCE components are located. These connections provide for redundancy as well as making use of basic SRST functionality in the event of a complete network failure, so that the remote site would still have basic dial tone service to make emergency (911) calls.

If side A of the WAN at the Unified CCE Agent Site fails, the following conditions apply:

- Any IP phones that are homed to the side-A Unified CM subscribers will automatically re-home to the side-B subscribers (provided the redundancy group is configured).
- Agent desktops that are connected to the CTI OS or Cisco Agent Desktop server at that site will automatically realign to the redundant CTI OS server at the remote site. (Agent desktops will be disabled during the realignment process.)

If both sides of the WAN at the Unified CCE Agent Site fail, the following conditions apply:

- The local voice gateway will detect the failure of the communications path to the Unified CM cluster and will go into SRST mode to provide local dial-tone functionality. With Unified CVP, these gateways detect the loss of the Unified CVP Call Server and execute their local survivability TCL script to reroute the inbound calls. Active calls in Unified CVP locally would no longer be visible to Unified CCE, so a Termination Call Detail (TCD) record would be written to the Unified CCE database at the time of the failure, and tracking of the call would stop at that point. The call would execute the local survivability TCL script, which could redirect it using the PSTN to another Unified CCE site that remains active; however, the call would then appear as a "new call" to Unified CCE and would have no relationship with the original call information. If the call is retained locally and redirected via SRST to a local phone, Unified CCE would not have visibility to the call from that point forward.
- The agent desktop will detect the loss of connectivity to the CTI OS Server (or Cisco Agent Desktop Server) and automatically log the agent out of the system. While the IP phones are in SRST mode, they will not be able to function as Unified CCE agents.

Understanding Failure Recovery

This section analyzes the failover recovery of each individual part (products and subcomponents inside each product) of the Unified CCE solution.

Unified CM Service

In larger deployments, it is possible that the Unified CM to which the agent phones are registered will not be running the CTI Manager service that communicates with the Unified CM Peripheral Gateway for Unified CCE. When an active Unified CM (call processing) service fails, all the devices registered to it are reported "out of service" by the CTI Manager service locally and to any external client, such as the Peripheral Gateway on a different subscriber CTI Manager service.

Unified CM call detail reporting (CDR) shows the call as terminated when the Unified CM failure occurred, although the call may have continued for several minutes after the failure because calls in progress stay in progress. IP phones of agents not on calls at the time of failure will quickly register with the backup Unified CM subscriber. The IP phone of an agent on a call at the time of failure will not register with the backup Unified CM subscriber until after the agent completes the current call. If MGCP, H.323, or SIP gateways are used, then the calls in progress survive, but further call control functions (hold, retrieve, transfer, conference, and so on) are not possible.

Unified CCE will also write a call record to the Termination Call Detail (TCD) table because Unified CM has reported the call as terminated to the Unified CCE PG. If the call continues after the PG has failed-over, a second TCD record will be written as a "new call" not related to the original call.

When the active Unified CM subscriber fails, the PG receives out-of-service events from Unified CM and logs out the agents. To continue receiving calls, the agents must wait for their phones to re-register with a backup Unified CM subscriber, then log back into their Unified CCE desktop application to have its functionality restored. Upon recovery of the primary Unified CM subscriber, the agent phones re-register to their original subscriber to return the cluster to the normal state, with phones and devices properly balanced across multiple active subscribers.

In summary, the Unified CM call processing service is separate from the CTI Manager service, which connects to the Unified CM PG via JTAPI. The Unified CM call processing service is responsible for registering the IP phones, and its failure does not affect the Unified CM PGs. From a Cisco Unified CCE perspective, the PG does not go off-line because the Unified CM server running CTI Manager remains operational. Therefore, the PG does not need to fail-over.

Unified IP IVR

When a CTI Manager service fails, the Unified IP IVR JTAPI subsystem shuts down and restarts by trying to connect to the secondary CTI Manager service on a backup Unified CM subscriber in the cluster. In addition, all voice calls at this Unified IP IVR are dropped. If there is an available secondary CTI Manager service on a backup subscriber, the Unified IP IVR logs into this CTI Manager service on that subscriber and re-registers all the CTI ports associated with the Unified IP IVR JTAPI user. After all the Unified CM devices are successfully registered with the Unified IP IVR JTAPI user, the server resumes its Voice Response Unit (VRU) functions and handles new calls. This action does not impact the Unified CVP because it does not depend upon the Unified CM CTI Manager service for call control.

Unified IP IVR Release 3.5 provided for cold standby and Release 4.0 provides hot standby redundancy, but this configuration is not supported for use with Unified CCE. These designs make use of a redundant server that is not used unless there is a failure of the primary Unified IP IVR server. However, during this failover processing, all calls that are in queue or treatment are dropped on the Unified IP IVR as part of the failover. A more resilient design would be to deploy a second (or more) Unified IP IVR server(s)

and have them all active, allowing Unified CCE to load-balance calls across them automatically. As shown in Figure 3-19, if one of the Unified IP IVR servers fails, only the calls on that server would fail, but the other active servers would remain active and be able to accept new calls in the system.

Unified CCE

Unified CCE is a collection of services and processes running on Unified CCE servers. The failover and recovery process for each of these services is unique and requires carefully examination to understand the impact to other parts of the Unified CCE solution, including another Unified CCE service.

Unified CM PG and CTI Manager Service

When the active CTI Manager Service or PG software fails, the PG JTAPI Gateway/PIM detects an OUT_OF_SERVICE event and induces a failover to the redundant (duplex) PG. Because the redundant PG is logged into the backup Unified CM subscriber CTI Manager Service already, it registers the IP phones and configured dialed numbers or CTI route points automatically. This initialization service takes place at a rate of about 5 devices per second. The agent desktops show them as being logged out or not ready, and a message displays stating that their routing client or peripheral (Unified CM) has gone off-line. (This warning can be turned on or off, depending on the administrator's preference.) All agents and supervisors lose their desktop third-party call control functionality until the failure recovery is complete. The agents and supervisors can recognize this event because call control action buttons on the desktop will gray out, and they will not be able to do anything with the desktop. Any existing calls remain active without any impact to the caller.

In the event that calls arrive at the CTI Route Points in Unified CM during a PG failover and the PIM is not yet fully operational, these calls will fail unless these route points are configured with a recovery number in their "Call Forward on Unregistered" or "Call Forward on Failure" setting. These recovery numbers could be the Cisco Unity voicemail system for the Auto Attendant, or perhaps the company operator position, to ensure the incoming calls are getting answered.

Note **Do** not push any buttons during desktop failover because these keystrokes can be buffered and sent to the CTI server when it completes its failover and restores the agent states.

When an active PG fails over to the idle side, calls still in progress will be recovered by querying Unified CM as part of the activation sequence. There will be one Termination Call Detail record providing information on the call, after the PG transition, when the call terminates. Peripheral call variables and ECC variables will be maintained on the agent desk top. Indication of whether the call was a barge-in or a conference call will be lost on the agent desktop and in reports, and an Intercept of a barged in call in progress will not be possible. Calls that were in the wrap-up state will be terminated. Agents will be able to release, transfer, or conference calls from their agent desktop after activation completes. During conference tear down, a call appearance from the desk top of an active call, but agent state will not be affected. Calls that end while the PG is down will be cut after a dead call time out after two hours.

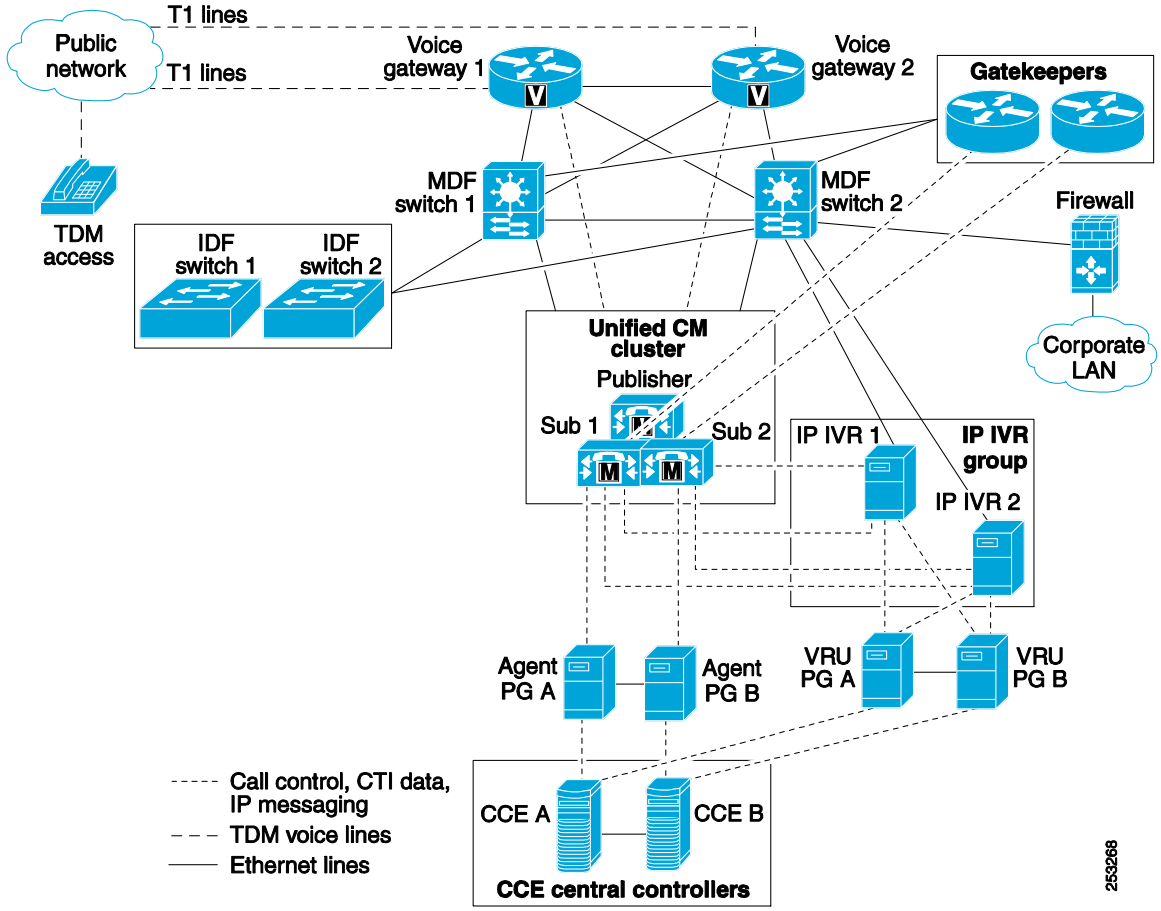
Note Call and agent state information might not be complete at the end of a failover if there are call status and agent state changes during the failover window.

Unified CCE Voice Response Unit PG

When a Voice Response Unit (VRU) PG fails, all the calls currently in queue or treatment on that Unified IP IVR are dropped unless there is a default script application defined or the CTI Ports have a recovery number defined in Unified CM for their "Call Forward on Failure" setting. Calls in progress or queued in Unified CVP are not dropped and will be redirected to a secondary Unified CVP or number in the H.323 or SIP dial plan, if available by the Survivability TCL script in the voice gateway.

The redundant (duplex) VRU PG side will connect to the Unified IP IVR or Unified CVP and begin processing new calls upon failover. Upon recovery of the failed VRU PG side, the currently running VRU PG continues to operate as the active VRU PG. Therefore, having redundant VRU PGs adds significant value because it allows a Unified IP IVR or Unified CVP to continue to function as an active queue point or to provide call treatment. Without VRU PG redundancy, a VRU PG failure would block use of that IP IVR even though the IP IVR is working properly. (See Figure 3-19.)

Figure 3-19 Redundant Unified CCE VRU PGs with Two IP IVR Servers



Unified CCE Call Router and Logger

The Unified CCE Central Controllers or Unified CCE Servers are shown in these diagrams as a single set of redundant servers. However, depending upon the size of the implementation, they could be deployed with multiple servers to host the following key software processes:

- Unified CCE Call Router

The Unified CCE Call Router is the brain of the system, and it maintains a constant memory image of the state of all the agents, calls, and events in the system. It performs the call routing in the system, executing the user-created Unified CCE Routing Scripts and populating the real-time reporting feeds for the Administration & Data Server. The Call Router software runs in synchronized execution, with both of the redundant servers running the same memory image of the current state across the system. They keep this information updated by passing the state events between the servers on the private LAN connection.

- Unified CCE Logger and Database Server

  The Unified CCE Logger and Database Server maintain the system database for the configuration (agent IDs, skill groups, call types, and so forth) and scripting (call flow scripts) as well as the historical data from call processing. The Loggers receive data from their local Call Router process to store in the system database. Because the Call Routers are synchronized, the Logger data is also synchronized. In the event that the two Logger databases are out of synchronization, they can be resynchronized manually by using the Unified ICMDBA application over the private LAN. The Logger also provides a replication of its historical data to the customer Administration & Data Server over the visible network.

In the event that one of the Unified CCE Call Routers fails, the surviving server will detect the failure after missing five consecutive TCP keep-alive messages on the private LAN. The Call Routers generate these TCP keep-alive messages every 100 ms, so it will take up to 500 ms to detect this failure. Upon detection of the failure, the surviving Call Router will contact the Peripheral Gateways in the system to verify the type of failure that occurred. The loss of TCP keep-alive messages on the private network could be caused by either of the following conditions:

- Private network outage — It is possible for the private LAN switch or WAN to be down but for both of the Unified CCE Call Routers to still be fully operational. In this case, the Peripheral Gateways will still see both of the Unified CCE Call Routers even though they cannot see each other over the private network to provide synchronization data. If the disabled synchronizer (Call Router B) can communicate with a majority of the PGs, it will then send a Test Other Side (TOS) message to the PGs sequentially to determine if the Call Router on the other side (Side A) is enabled. If Call Router B receives a message that side A is in fact enabled, then Call Router A will run in simplex until the private network is restored. If all the PGs reply to the TOS message and indicate that side A is down, then side B re-initializes in simplex mode.
- Call Router hardware failure — It is possible for the Call Router on the other side to have a physical hardware failure and be completely out of service. In this case, the Peripheral Gateways would report that they can no longer see the Call Router on the other side, and the surviving Call Router would take over the active processing role in simplex mode. This failure is detected by the Call Routers from the loss of heartbeat keep-alives on the Private Network.

During the Call Router failover processing, any Route Requests sent to the Call Router from a Carrier Network Interface Controller (NIC) or Peripheral Gateway will be queued until the surviving Call Router is in active simplex mode. Any calls in progress in the IVR or at an agent will not be impacted.

If one of the Unified CCE Logger and Database Servers were to fail, there would be no immediate impact except that the local Call Router would no longer be able to store data from call processing. The redundant Logger would continue to accept data from its local Call Router. When the Logger server is restored, the Logger will contact the redundant Logger to determine how long it had been off-line. If the Logger was off-line for less than 12 hours, it will automatically request all the transactions it missed from the redundant Logger while it was off-line. The Loggers maintain a recovery key that tracks the date and time of each entry recorded in the database, and these keys are used to restore data to the failed Logger over the private network.

If the Logger was off-line for more than 12 hours, the system will not automatically resynchronize the databases. In this case, resynchronization has to be done manually using the Unified ICMDBA application. Manual resynchronization allows the system administrator to decide when to perform this data transfer on the private network, perhaps scheduling it during a maintenance window when there would be little call processing activity in the system.

The Logger replication process that sends data from the Logger database to the HDS database on the Administration & Data Servers will automatically replicate each new row written to the Logger database when the synchronization takes place as well.

There is no impact to call processing during a Logger failure; however, the historical data on the Administration & Data Server that is replicated from that Logger would stop until the Logger can be restored.

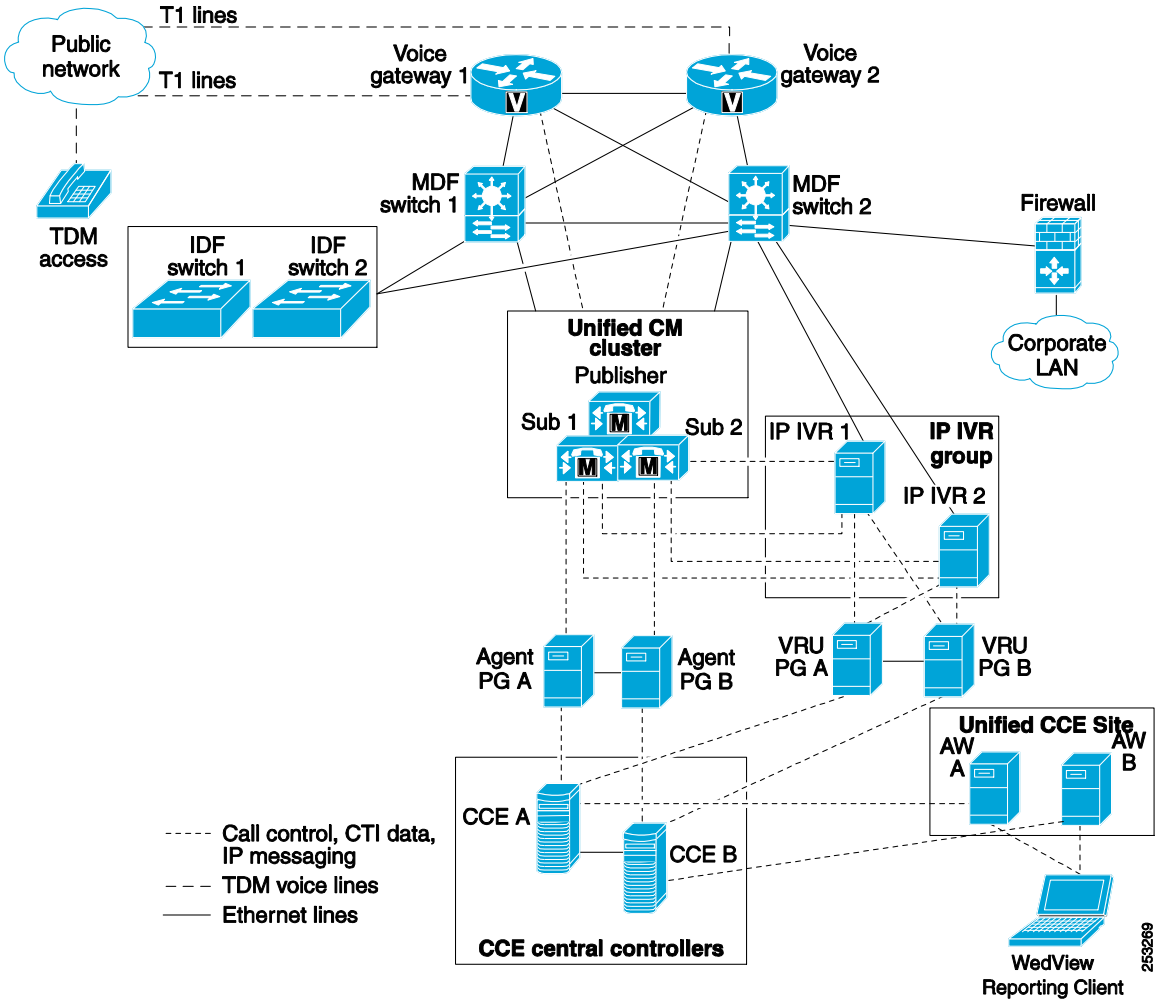
Additionally, if the Unified Outbound Option is used, the Campaign Manager software is loaded on Logger A only. If that platform is out of service, any outbound calling will stop until the Logger can be restored to operational status.

Administration & Data Server

The Administration & Data Server provides the user interface to the system for making configuration and scripting changes. It also can host the web-based reporting tool, WebView and Internet Script Editor.

These servers do not support redundant or duplex operation, as the other Unified CCE system components do. However, you can deploy multiple Administration & Data Servers to provide redundancy for Unified CCE. (See Figure 3-20.)

Figure 3-20 Redundant Unified CCE Administration & Data Servers



Administration & Data Server Real-Time Distributors are clients of the Unified CCE Call Router real-time feed that provides real-time information about the entire Unified CCE across the enterprise. Real-Time Distributors at the same site can be set up as part of an Admin Site that includes a designated primary real-time distributor and one or more secondary real-time distributors. Another option is to add Administration Clients that do not have their own local SQL databases and are homed to a Real-Time Distributor locally for their SQL database and real-time feed.

The Admin Site reduces the number of real-time feed clients the Unified CCE Call Router has to service at a particular site. For remote sites, this is important because it can reduce the required bandwidth to support remote Administration & Data Servers across a WAN connection.

When using an Admin Site, the primary Administration & Data Server is the one that will register with the Unified CCE Call Router for the real-time feed, and the other Administration & Data Servers within that Admin Site register with the primary Administration & Data Server for the real-time feed. If the primary real-time distributor is down, the secondary real-time distributors will register with the Unified CCE Call Router for the real-time feed. Administration Clients that cannot register with the primary or secondary Administration & Data Server, will not be able to perform any tasks until the distributors are restored.

Alternatively, each Administration & Data Server could be deployed in its own Admin Site regardless of the physical site of the device. This deployment will create more overhead for the Unified CCE Call Router to maintain multiple real-time feed clients; however, it will prevent a failure of the primary Administration & Data Server from taking down the secondary Administration & Data Server at the site.

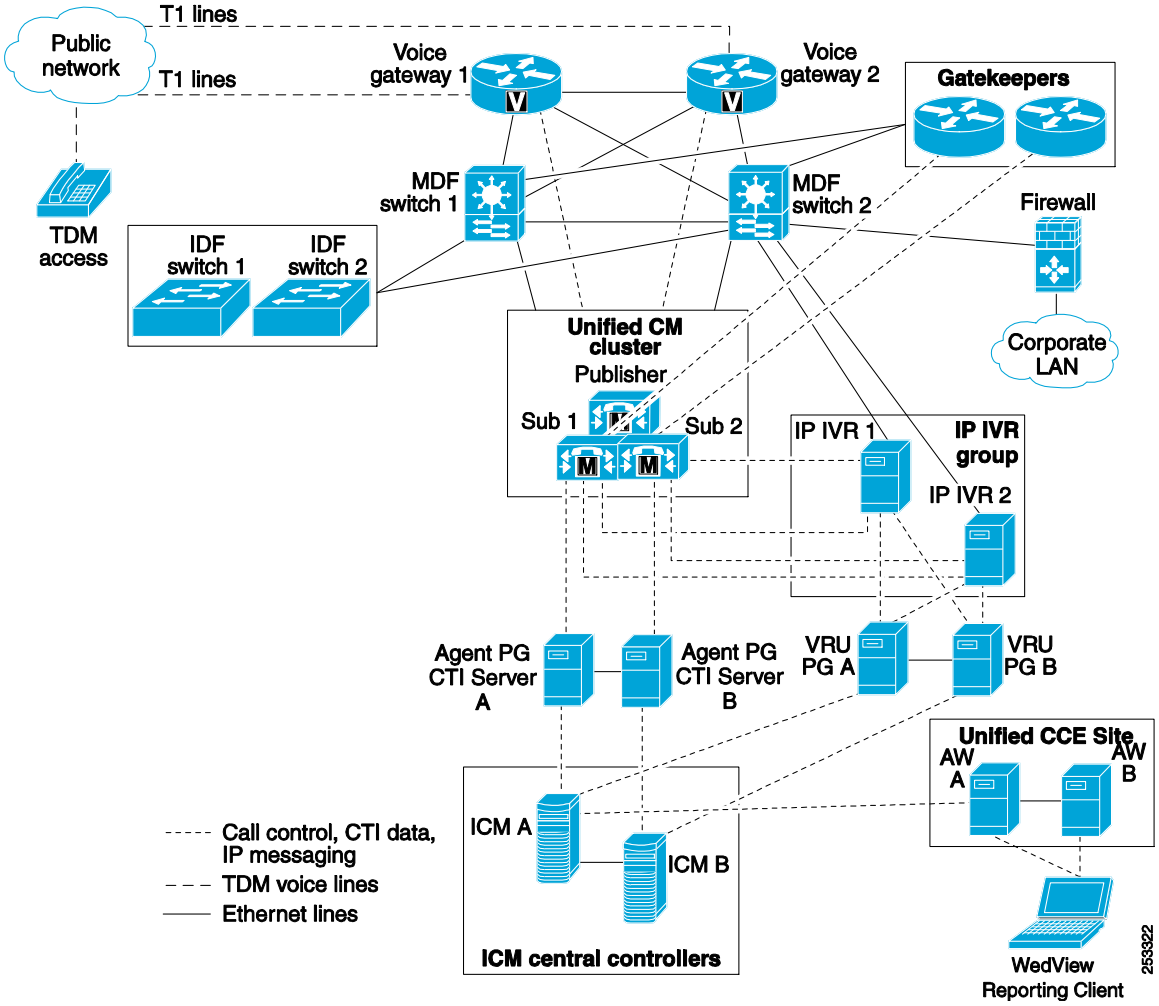
Additionally, if the Administration & Data Server is being used to host the ConAPI interface for the Cisco Unified Contact Center Management Portal (Unified CCMP), any configuration changes made to the Unified CCE or Unified CCMP systems will not be passed over the ConAPI interface until it is restored.

CTI Server

The CTI Server monitors the data traffic of the Unified CM PIM on the Agent PG for specific CTI messages (such as call ringing or off-hook events) and makes those messages available to CTI clients such as the CTI OS Server or Cisco Agent Desktop Enterprise Server. It also processes third-party call control messages (such as make call or answer call) from the CTI clients and sends those messages via the PIM interface of the PG to Unified CM to process the event on behalf of the agent desktop.

CTI Server is redundant and co-resident on the Agent PG servers. (See Figure 3-21.) It does not, however, maintain agent state in the event of a failure. Upon failure of the CTI Server, the redundant CTI server becomes active and begins processing call events. CTI OS Server is a client of the CTI Server and is designed to monitor both CTI Servers in a duplex environment and maintain the agent state during failover processing. CTI OS agents will see their desktop buttons gray-out during the failover to prevent them from attempting to perform tasks while the CTI Server is down. The buttons will be restored as soon as the redundant CTI Server is restored, and the agent does not have to log on again to the desktop application. Most call context will be maintained, but ANI and DNIS will be lost in this instance where only the CTI Server component is impacted.

Figure 3-21 Redundant CTI Servers Co-Resident on Agent PG

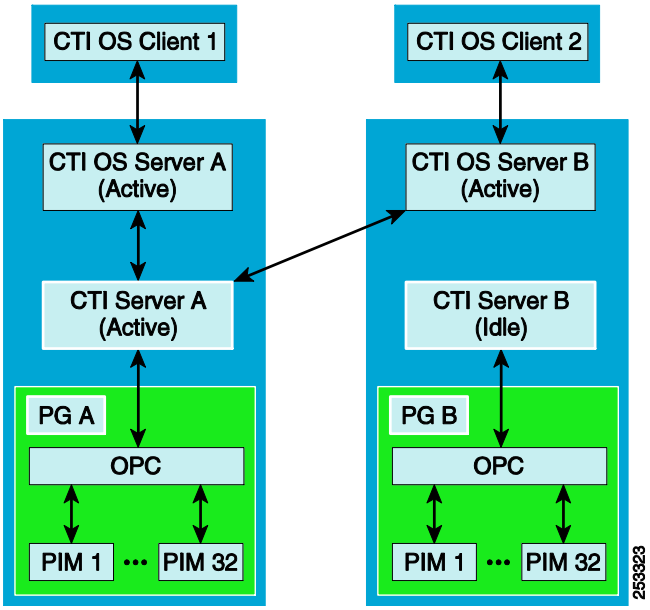


CTI OS Considerations

CTI OS Server is a software component that runs co-resident on the Unified CM Peripheral Gateway. CTI OS Server software is designed to be fault-tolerant and is typically deployed on redundant physical servers; however, unlike the PG processes that run in hot-standby mode, both of the CTI OS Server processes run in active mode all the time. The CTI OS Server processes are managed by NodeManager, which monitors each process running as part of the CTI OS service and which automatically restarts abnormally terminated processes.

CTI OS handles failover of related components as described in the following scenarios (see Figure 3-22).

Figure 3-22 Redundant CTI OS Server Processes



Scenario 1: CTI Server Side A (Active) Fails

In this scenario, CTI Server side A is co-resident on PG side A, and the following events occur:

- CTI Server side B detects the failure of side A and becomes active.
- NodeManager restarts CTI Server side A and becomes idle.
- Both CTI OS Server sides A and B drop all CTI OS client/agent connections and restart after losing the connection to CTI Server A. At startup, CTI OS Server sides A and B stay in CONNECTING state until they connect to CTI Server side B, and then they go into CONFIGURING state, where they download agent and call states and configuration information. CTI OS Client connections are not accepted by CTI OS Server A and B during CONNECTING and CONFIGURING states. When CTI OS Server synchronizes with CTI Server, the state becomes ACTIVE and it is now ready to accept CTI OS Client connections.
- Both CTI OS Clients 1 and 2 lose connections to CTI OS Servers, and they each randomly select one CTI OS Server to connect to. CTI OS Client 1 can be connected to either CTI OS Server A or B, and the same is true for CTI OS Client 2. During this transition, the buttons of the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to a CTI OS.

Scenario 2: CTI Server B (Idle) Fails

In this scenario, CTI Server side B is co-resident on PG side B but is not the active side. The following events occur:

- CTI Server side A stays active.
- NodeManager restarts CTI Server side B and stays idle.
- Neither CTI OS Clients nor CTI OS Servers are affected by this failure.

Scenario 3: CTI OS Server A Fails

In this scenario, CTI OS Server side A processes are co-resident on PG/CTI Server side A. The following events occur:

- CTI OS Client 1 detects the loss of network connection and automatically connects to CTI OS server B. During this transition, the buttons of the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to CTI OS server B.
- CTI OS Client 2 stays connected to CTI OS Server B.
- NodeManager restarts CTI OS Server A.

Scenario 4: CTI OS Server B Fails

In this scenario, CTI OS Server side A processes are co-resident on PG/CTI Server side B. The following events occur:

- CTI OS Client 2 detects the loss of network connection and automatically connects to CTI OS server A. During this transition, the buttons of the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to CTI OS server A.
- CTI OS Client 1 stays connected to CTI OS Server A.
- NodeManager restarts CTI OS Server B.

Scenario 5: CTI OS Client 1 Fails

In this scenario, the following events occur:

- The agent manually restarts CTI OS Client 1 application.
- CTI OS Client 1 randomly selects one CTI OS Server to connect to. (CTI OS Client 1 can be connected to either CTI OS Server A or B.)
- Once connected, the agent logs in, and CTI OS Client 1 recovers its state by getting agent and call states through the CTI OS Server to which it is connected.

Scenario 6: CTI OS Client 2 Fails

In this scenario, the following events occur:

- The agent manually restarts CTI OS Client 2 application.
- CTI OS Client 2 randomly selects one CTI OS Server to connect to. (CTI OS Client 2 can be connected to either CTI OS Server A or B.)
- Once connected, the agent logs in, and CTI OS Client 2 recovers its state by getting agent and call states through the CTI OS Server to which it is connected.

Scenario 7 - Network Failure Between CTI OS Client 1 and CTI OS Server A

In this scenario, the following events occur:

- CTI OS Server A drops the connection of CTI OS Client 1
- CTI OS Client 1 detects the loss of network connection and automatically connects to CTI OS server B. During this transition, the buttons of the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to CTI OS server B.

Scenario 8: Network Failure Between CTI OS Client 1 and CTI OS Server B

CTI OS Client 1 is not affected by this failure because it is connected to CTI OS Server A.

Scenario 9: Network Failure Between CTI OS Client 2 and CTI OS Server A

CTI OS Client 2 is not affected by this failure because it is connected to CTI OS Server B.

Scenario 10: Network Failure Between CTI OS Client 2 and CTI OS Server B

In this scenario, the following events occur:

- CTI OS Server B drops the connection of CTI OS Client 2.
- CTI OS Client 2 detects the loss of network connection and automatically connects to CTI OS server A. During this transition, the buttons of the CTI Toolkit Agent Desktop will be disabled and will return to operational state as soon as it is connected to CTI OS server A.

Cisco Agent Desktop Considerations

Cisco Agent Desktop

Cisco Agent Desktop client applications are a client of CTI OS, which provides for automatic failover and redundancy for the Cisco Agent Desktop CTI connections. If the Unified CM Peripheral Gateway or CTI Server (CG) fails-over, the Cisco Agent Desktop client application will display a logged out state and will automatically return to a logged in state when an operational connection is established with the alternate Unified CM Peripheral Gateway or CTI Server (CG). Consequently the scenarios outlined in CTI OS Considerations section apply.

The Cisco Agent Desktop services (Enterprise, Chat, RASCAL, and so forth) can also be deployed redundantly to allow for failover of the core Cisco Agent Desktop components. The Cisco Agent Desktop client applications are aware of the redundant Cisco Agent Desktop services and will automatically failover in the event of a Cisco Agent Desktop service process or hardware failure.

The following services are active on only one side at a time:

- Cisco Browser and IP Phone Agent Service
- Cisco Chat Service
- Cisco Enterprise Service
- Cisco Licensing and Resource Manager Service
- Cisco Recording and Statistics Service
- Cisco Sync Service

The following services are active on both sides at all times and will be available to the CAD client applications as long as network connectivity is available:

- Cisco LDAP Monitor Service
- Cisco Recording & Playback Service
- Cisco VoIP Monitor Service

Scenario 1: Cisco Agent Desktop License and Resource Manager Service Side A (Active) Fails

In this scenario, the following events occur:

- The CAD services on side A that are not always active will go to an idle state.
- The CAD services on side B (idle) will activate.
- The CAD client applications will recover to side B.

Scenario 2: A Cisco Agent Desktop Service on Side A (Active) Fails Twice Within Five Minutes

In this scenario, the following events occur:

- The CAD services on side A that are not always active will go to an idle state.

- The CAD services on side B (idle) will activate.
- The CAD client applications will recover to side B.

Scenario 3: A Cisco Agent Desktop Service on Side A (Active) Fails and Remains Down for Three Minutes

In this scenario, the following events occur:

- The CAD services on side A that are not always active will go to an idle state.
- The CAD services on side B (idle) will activate.
- The CAD client applications will recover to side B.

Scenario 4: Network Failure Between CAD Services on Side A (Active) and CAD Services on Side B (Idle)

In this scenario, the following events occur:

- The CAD services on side A will remain active.
- The CAD services on side B (idle) will activate.
- The CAD client applications will remain connected to CAD services on side A.
- When network connectivity is restored between sides A and B the Cisco Licensing and Resource Manager Service will render inactive the non-preferred side. Recovery side preference is configurable in Post Install.

Cisco Agent Desktop Browser Edition and IP Phone Agent

Cisco Agent Desktop Browser Edition and IP Phone Agent communicate with CTI Server via the Cisco Browser and IP Phone Agent service. When launching CAD-BE, the agent may use the URL for either side as long as the desired side is accessible. Once launched, CAD-BE will automatically connect to the active side. When launching IPPA, the agent must select the active side from the services menu on the phone. If the idle side is selected, the user will receive an error informing them that the side selected is idle and to try the other side.

Scenario 1: Cisco Agent Desktop Services on Side A (Active) Fail and Side B (Idle) Becomes Active

In this scenario, the following events occur for a logged-in CAD-BE agent:

- The CAD-BE applet will change to a logged out state and the user will be notified that the connection has been lost.
- The CAD-BE applet will automatically connect to services on side B and log-in the agent.

In this scenario, the following events occur for a logged-in IPPA agent:

- The IPPA agent will be notified that the connection to the server has been lost.
- The IPPA agent must manually select side B from their services list and log in again.

Replacement of MSDE with Flat Files

As MSDE is no longer supported, at post install time the user can choose flat files or a full SQL database. Postinstall will configure their system based on their selection. There will be a value stored in LDAP that indicates which implementation is selected. Once the implementation is selected (and saved, when the initial configuration of postinstall is completed and saved), the user cannot change implementations. For the database implementation, Unified CCE configures the FCRasSvr database, as it does now. Unified CCE will continue to provide scripts for setup and teardown of database replication for HA. In

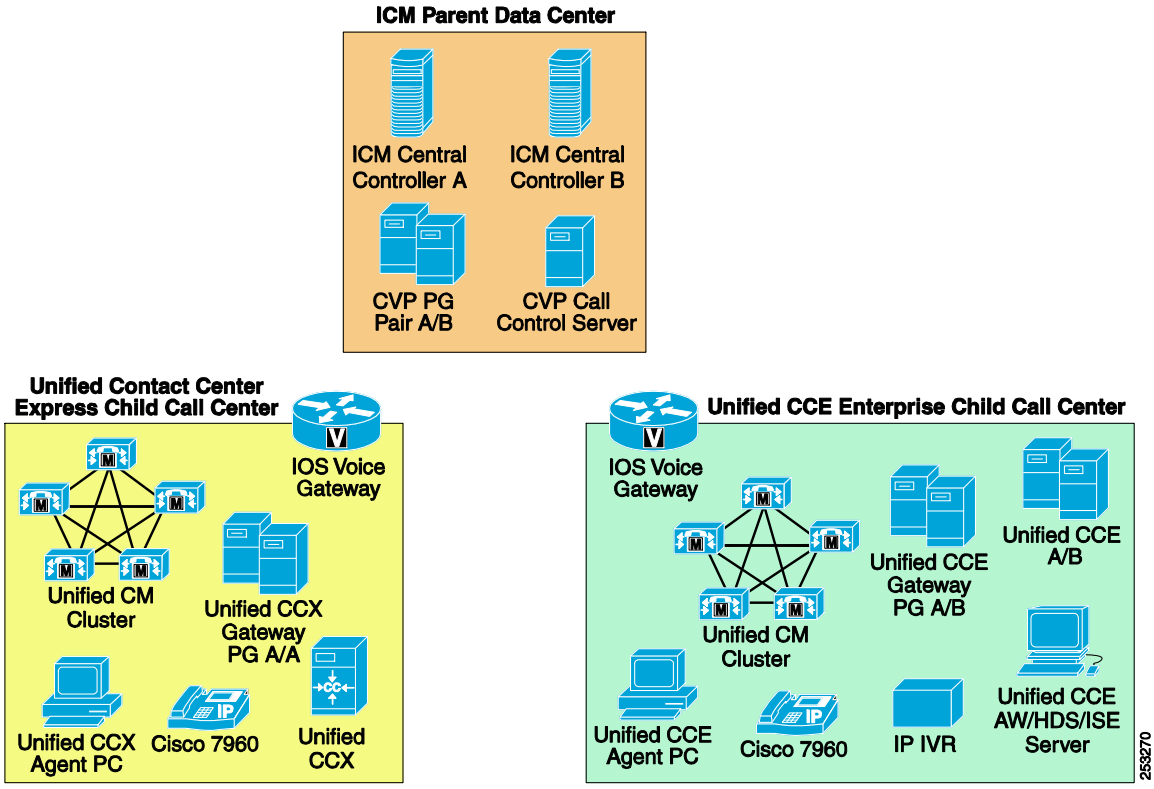
the database implementation, there are three tables: agent state data, call log data, and recording metadata. In the flat-file implementation, each of these tables is represented by a set of text files. Regardless of which implementation is used in a particular installation, the data stored will be identical.

There are currently no minimum or maximum file sizes set.

Design Considerations for Unified CCE Deployment with Unified ICM Enterprise

The parent/child deployment is where the Unified ICM acts as the parent controlling one or more Unified CCE child IP ACDs. (See Figure 3-23.) In this model, the Unified ICM Enterprise system is designed to be the network call routing engine for the contact centers, with network queuing using the Unified CVP and Unified CCE Gateway Peripheral Gateways to connect child Unified CCE systems (either Unified CCE with system PG or Unified CCX). The child Unified CCE systems are individual IP-ACD systems, fully functional with local call processing in case they lose their WAN connection to the parent Unified ICM system. This configuration provides a high level of redundancy and availability to the Unified CCE solution to allow sites to remain functional as Unified CCE sites even if they are cut off from centralized call processing resources.

Figure 3-23 Parent/Child Deployment Model



Parent/Child Components

The following sections describe the components used in Unified ICM Enterprise (parent) and Unified CCE System (child) deployments.

The Unified ICM Enterprise (Parent) Data Center

The Unified ICM parent data center location contains the Unified ICM Central Controller. In Figure 3-23, it is shown as a redundant (duplex) pair of Central Controllers, which represents Call Router and Logger servers. These servers can be deployed as individual Call Routers and Loggers, and they can also be deployed in two different data centers to be geographically distributed for additional fault tolerance.

The Unified ICM Central controllers control Peripheral Gateways at the data center location. In Figure 3-23, there is only a redundant (duplex) pair of IVR PGs used to control Unified CVP across the architecture. Additional PGs can be inserted at this layer to control TDM or legacy ACDs and IVRs, perhaps to support a migration to Unified CCE or to support out-source locations that still use the TDM or legacy ACDs. The Unified ICM parent at this level can also support standard pre-routing with inter-exchange carriers (IXCs) such as AT&T, Sprint, MCI, and others, thus allowing Unified ICM to select the best target for the call while it is still in the carrier network.

The Unified ICM parent is not designed to support any directly controlled agents in this model, which means that it does not support classic Unified CCE with a Unified CM Peripheral Gateway installed on this Unified ICM parent. All agents must be controlled externally to this Unified ICM parent system.

The Unified CVP or IVR PG pair controls the Customer Voice Portal Call Server, which translates the IVR PG commands from Unified ICM into VoiceXML and directs the VoiceXML to the voice gateways at the remote contact center sites. This allows calls from the data center location to come into the remote call centers under control of the Unified CVP at the parent location. The parent then has control over the entire network queue of calls across all sites and will hold the calls in queue on the voice gateways at the sites until an agent becomes available.

The Unified Contact Center Express (CCX) Call Center (Child) Site

The Unified Contact Center Express (CCX) Call Center location contains a local Unified CM cluster that provides local IP-PBX functionality and call control for the IP phones and local Unified CVP voice gateway. There is also a local Unified CCX Server that provides IP-ACD functionality for the site. Prior to Unified CCX Server Release 8.0, the Unified CCX Server had the Unified CCE Gateway PG installed on it, which reduces the number of servers required to support this contact center site. Unified CCX 8.0(1) is deployed on the Unified Communications Operating System platform, requiring the Unified CCE Gateway PG to be installed on a separate (Windows) server. The deployment model changes for new and existing customers require the Unified CCE Gateway PG and the CCX ACMI Manager to be installed on separate (Windows) servers. In either of these deployments, the Unified CCE Gateway PG connects to the Unified ICM Call Router (Rogger) at the Unified ICM parent data center location over the WAN and provides real-time event data and agent states to the parent from the Unified CCX. The Unified CCE Gateway PG also captures configuration data (skill groups, CSQs, services, applications, and so forth) and sends it to the parent Unified ICM configuration database as well.

Additional Unified CCX servers may be used and included in this site to provide redundant Unified CCX Servers, historical reporting database services, recording and monitoring servers, and ASR/TTS servers as well. Highly available deployments of Unified CCX Release 8.0 or above, requires the deployment of two "SideA" Unified CCE Gateway PGs on separate (Windows) servers. The Unified CCX Server(s) are configured with the IP Addresses of the two "SideA" Unified CCE Gateway PGs.

The Unified CCE Call Center (Child) Site

The Unified CCE Call Center location contains a local Unified CM cluster that provides local IP-PBX functionality and call control for the IP phones and local Unified CVP voice gateway. There is also a local Unified IP IVR to provide local call queuing for the Unified CCE site. There is a redundant pair of

Unified CCE Gateway PGs that are used to connect this site to the Unified ICM parent Central Controller at the Unified ICM parent data center location over the WAN. The Unified CCE Gateway PGs can be deployed on separate servers or co-resident with the CCE System PG with the following caveats:

If the Unified CCE Gateway PG and Unified CCE System PG Instance Numbers are the same, then the PG number for the Unified CCE Gateway PG and Unified CCE System PG must be different.

If the Unified CCE Gateway PG and Unified CCE System PG Instance Numbers are different, then the PG number for Unified CCE Gateway PG and Unified CCE System PG may be the same.

No additional PGs (such as VRU PG or MR PG) can be added to this Server.

For scalability limits of the co-resident Unified CCE Gateway PG and Unified CCE System PG, refer to Sizing Unified CCE Components and Servers, page 10-1 for additional details.

The Unified CCE Gateway PGs provide real-time event data and agent states to the parent from the Unified CCE child. The Unified CCE Gateway PGs also capture configuration data (skill groups, services, call types, and so forth) and send it to the parent Unified ICM configuration database as well.

The IP-IVR at the Child site can be replaced with a local Unified CVP instance. Unified CVP is not integrated as part of the Agent Controller's System PG; there is a separate IVR PG defined specifically for Unified CVP as part of the installation for System CCE with Unified CVP. Because Unified CVP is not part of the System PG, calls in queue or treatment in Unified CVP will not be reported to the Parent ICM via the Unified CCE Gateway PG.

A local Unified CCE child system is used to provide IP-ACD functionality, and it can be sized depending upon the type of deployment required:

- Progger configuration

  Single (or duplex) server that contains the Unified CCE components: Call Router and Logger, System PG for Unified CM and IP IVR, CTI Server and CTI OS Server, and optionally the VRU PG for Unified CVP.

- Rogger configuration with separate Unified CCE Agent Controller (System PG and optional Unified CVP controller and CTI/CTI OS Server)

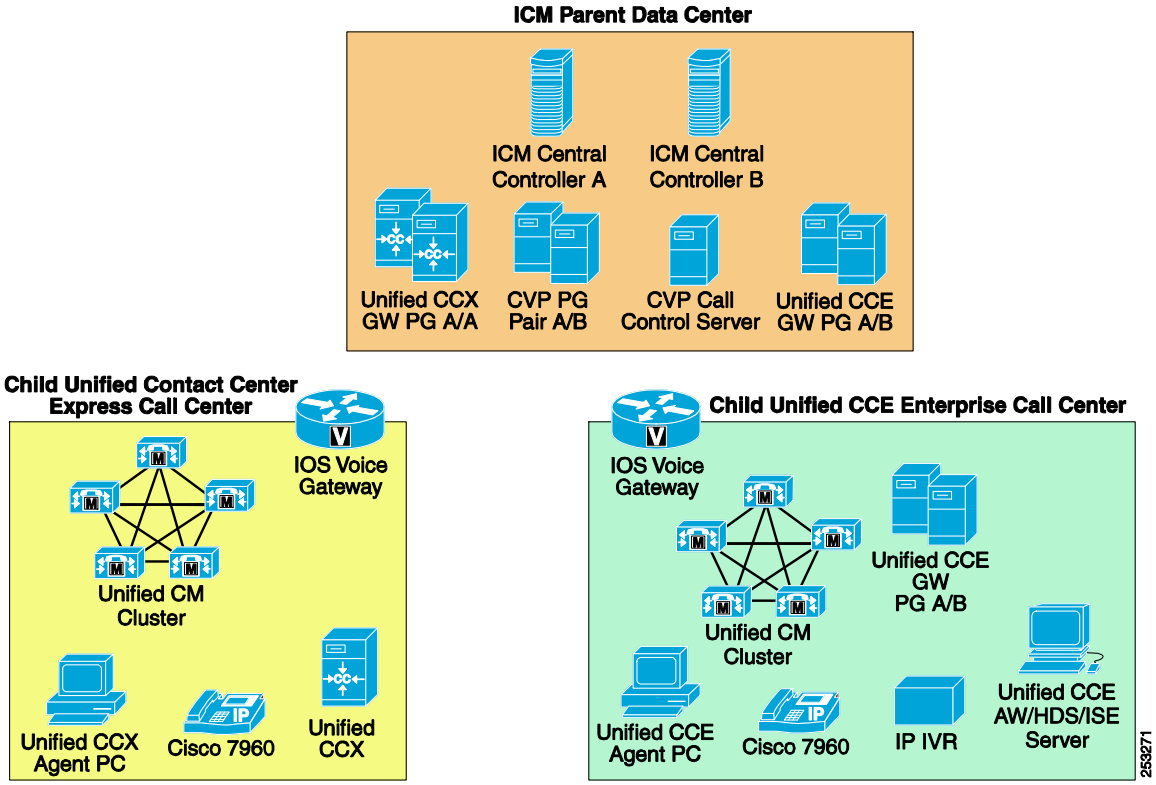
  The Rogger configuration contains the Unified CCE components: Call Router and Logger as a single set of duplex Central Controllers, and a separate Agent Controller set of duplex servers that contain the System PG for Unified CM and IP IVR, CTI Server and CTI OS Server, and the optional VRU PG for Unified CVP.

For more details about the capacity of these configurations, refer to Sizing Unified CCE Components and Servers, page 10-1.

In either configuration, a separate Administration & Data Server is required to host the configuration and scripting tools for the system as well as an optional Historical Database Server role and Web-based reporting (WebView) tool.

Unified ICM Enterprise (Parent) with Unified CCE Gateway PGs at Data Center

Figure 3-24 Parent/Child Deployment Model with Unified CCE Gateway PGs at Data Center



The Unified CCE Gateway PG may be deployed at the ICM Parent Data Center, as illustrated in Figure 3-24. Some advantages with this deployment model include centrally managing and controlling Unified CCE Gateway PGs as well as configuring Unified CCE Gateway PGs with multiple PIMs to help reduce the server TCO requirements. Another condition forcing the Unified CCE Gateway PG to be deployed at the ICM Parent Data Center is where the ownership and management of the Unified CCE and Parent Data Center are different. For example the Unified CCE is managed by an Outsourcer/Service Bureau, and the Parent manages the Unified CCE Gateway PG.

There are several drawbacks with moving the Unified CCE Gateway PGs to the Data Center. One is specific to recovering reporting data in the event of a network failure. If the network connection between the Parent Data Center and the Child Unified CCE System PGs drops, all reporting at the parent is lost for that period. NOTE: If the Unified CCE Gateway PG is deployed locally to the Unified CCE System PG, and the connection between the Unified CCE Gateway PG and the Parent Data center drops, the parent historical data is updated when the network connection is restored.

A second drawback with centralizing the Unified CCE Gateway PGs is that the network bandwidth requirements for the connections between the Parent CCE Gateway PG and the Child CCE System PG are significantly higher. Refer to the "Bandwidth Requirements for Unified CCE Gateway to System PG" section in the Bandwidth Provisioning and QoS Considerations chapter for additional details.

Parent/Child Call Flows

The following sections describe the call flows for the parent and child.

Typical Inbound PSTN Call Flow

In a typical inbound call flow from the PSTN, calls would be directed by the carrier network to the contact center sites using some predefined percent allocation or automatic routing method. These calls are terminated in the Unified CVP voice gateways at the call center locations, under control of the Unified ICM parent Unified CVP. The inbound call flow is as follows:

1. The call arrives on the Unified CVP voice gateway at the Unified CCE call center location.
2. The Unified CVP voice gateway maps the call by dialed number to a particular Unified CVP Call Server at the Unified ICM parent site and sends a new call event to the Unified CVP Call Server.
3. The Unified CVP Call Server sends the new call event message to the Unified CVP or IVR PG at the Unified ICM parent site.
4. The Unified CVP PG sends the new call message to the Unified ICM parent, which uses the inbound dialed number to qualify a routing script to determine the proper call treatment (messaging) or agent groups to consider for the call.
5. Unified ICM instructs Unified CVP to hold the call in the voice gateway at the site and wait for an available agent, while directing specific instructions to play .wav files for hold music to the caller in the gateway.
6. When an agent becomes available, the Unified ICM instructs Unified CVP to transfer the call to the site with the available agent by using a translation route. (The agent might not be at the same physical site but across the WAN.) Any data collected about the call in the Unified ICM parent Unified CVP will be transferred to the remote system's PG (either a TDM, legacy PG, or one of the Unified CCE Gateway PGs for Unified CCX or Unified CCE).
7. When the call arrives at the targeted site, it will arrive on a specific translation route DNIS that was selected by the Unified ICM parent. The PG at the site is expecting a call to arrive on this DNIS to match up with any pre-call CTI data associated with the call. The local ACD or Unified CCE will perform a post-route request to the local PG to request the CTI data as well as the final destination for the call (typically the lead number for the skill group of the available agent).
8. If the agent is no longer available for the call (walked away or unplugged), Unified CVP at the Parent site will use the Router Requery function in the ICM Call Routing Script to select another target for the call automatically.

Post-Route Call Flow

Post-routing is used when a call is already at a peripheral ACD or IVR and needs to be routed intelligently to another agent or location. If an agent gets a call in the ACD or Unified CCE that needs to be sent to a different skill group or location, the agent can make use of the post-route functionality to reroute the call. The post-route call flow is as follows:

1. The agent transfers the call to the local CTI route point for reroute treatment using the CTI agent desktop.
2. The reroute application or script makes a post-route request to the Unified ICM parent via the local Unified CCE Gateway PG connection.
3. The Unified ICM parent maps the CTI route point from Unified CCE as the dialed number and uses that number to select a routing script. This script will return a label or routing instruction that can move the call to another site, or to the same site but into a different skill group, or to a Unified CVP node for queueing.
4. The Unified CCE receives the post-route response from the Unified ICM parent system and uses the returned routing label as a transfer number to send the call to the next destination.

Parent/Child Fault Tolerance

The parent/child model provides for fault tolerance to maintain a complete IP-ACD with either Unified CCX or Unified CCE deployed at the site, with local IP-PBX and call treatment and queueing functionality.

Unified CCE Child Loses WAN Connection to Unified ICM Parent

If the WAN between the Unified CCE child site and the Unified ICM parent fails, the local Unified CCE system will be isolated from the parent as well as the Unified CVP voice gateway. Calls coming into the site will no longer get treatment from the Unified CVP under control of the Unified ICM parent, so the following functionality must be replicated locally, depending on the Child configuration.

- For Unified CCE Child configurations using local IP IVR resources for queue and treatment:
  - The local voice gateway must have dial peer statements to pass control of the calls to the local Unified CM cluster if the Parent Unified CVP Call Server cannot be reached. Also, the local Unified CM cluster must have CTI route points mapped to the inbound DNIS or dialed numbers that the local voice gateway will present if the Parent Unified CVP Call Server is not reached.
  - The local IP IVR must be configured with appropriate .wav files and applications that can be called by the Unified CCE Child system locally to provide basic call treatment such as playing a welcome greeting or other message.
  - The Child CCE Routing Script must handle queueing of calls for agents in local skill groups, instructing the IP IVR to play treatment in-queue while waiting for an agent.
  - Any data lookup or external CTI access that is normally provided by the Parent Unified CVP or the Parent Unified ICM must be provisioned locally to allow the agents to have full access to customer data for routing and screen pops.
  - Any post-routing transfer scripts will fail during this outage, so Unified CCE must be configured to handle this outage or prevent the post-route scripts from being accessed.
- For Unified CCE Child configurations using local Unified CVP resources for queue and treatment with Unified CCE 7.5*(x)*:
  - The local voice gateway must have dial peer statements to pass control of the calls to the local Unified CVP Call Server at the Child site. Also, the inbound DNIS or dialed numbers that the local voice gateway will present to the Child Unified CVP must be configured in the Child Unified CCE to process these calls locally at the Child.
  - The local VXML Gateways and Unified CVP Call Servers must be configured with appropriate .wav files and applications that can be called by the Unified CCE Child system locally to provide basic call treatment such as playing a welcome greeting or other messages.
  - Self-service or Unified CVP Studio VXML applications normally provided by the Parent Unified ICM must be replicated using Unified CVP VXML Server (web application server) at the Child site to generate the dynamic VXML for these applications.
  - The Child Unified CCE Routing Script must handle queueing of calls for agents in local skill groups, instructing the local Unified CVP at the Child site to play treatment in-queue while waiting for an agent.
  - Any data lookup or external CTI access that is normally provided by the Parent Unified CVP or the Parent Unified ICM must be provisioned locally to allow the agents to have full access to customer data for call routing and screen pops.
  - Any post-routing transfer scripts will fail during this outage, so Unified CCE must be configured to handle this outage or prevent the post-route scripts from being accessed.

Unified Contact Center Express Child Loses WAN Connection to Unified ICM Parent

If the WAN between the Unified Contact Center Express (CCX) child site and the Unified ICM parent fails, the local Unified CCX system will be isolated from the parent as well as the Unified CVP voice gateway. Calls coming into the site will no longer get treatment from the Unified CVP under control of the Unified ICM parent, so the following functionality must be replicated locally:

- The local voice gateway must have dial peer statements to pass control of the calls to the local Unified CM cluster if the Parent Unified CVP Call Server cannot be reached.
- Unified CCX JTAPI applications have to be mapped to these CTI route points to provide any typical inbound call treatment, such as playing a welcome greeting or other message.
- The application has to provide for call queueing and treatment in queue while waiting for a local Contact Service Queue (CSQ) agent.
- Any data lookup or external CTI access that is normally provided by the Parent Unified CVP or the Parent Unified ICM must be provisioned locally to allow the agents to have full access to customer data for call routing and screen pops.
- Any post-routing applications or transfer scripts will fail during this outage, so the Unified CCX must be configured to handle this outage or prevent the post-route applications from being accessed.

A similar failure would occur if the local Unified CVP ingress voice gateways controlled by the Parent Unified CVP Call Server could not see the Unified ICM Parent CVP Call Servers. The local Unified CVP gateways would be configured to fail-over to the local Unified CM (or Child Unified CVP) to route calls to the Unified CCX agents as described above. Likewise, if the entire Unified ICM parent were to fail, the local voice gateways controlled by the Parent Unified CVP at the sites would no longer have call control from the Unified ICM parent, and calls would forward to the local sites for processing.

Unified CCE Gateway PG Fails or Cannot Communicate with Unified ICM Parent

If the Unified CCE gateway PG fails or cannot communicate with the Unified ICM parent, the local agents are no longer seen as available to the Unified ICM parent, but the inbound calls to the site may still be under control of the Unified ICM parent Unified CVP. In this case, the Unified ICM parent will not know if the remote Unified CCE gateway PG has failed or if the actual Unified CCE IP-ACD has failed locally.

The Unified ICM at the parent location can automatically route around this site, considering it down until the PG comes back online and reports agent states again. Alternatively, the Unified ICM can also direct a percentage of calls as blind transfers to the site Unified CCE or Unified CCX using the local inbound CTI route points on Unified CM. This method would present calls with no CTI data from Unified CVP, but it would allow the agents at the site to continue to get calls locally with their Unified CCE/CCX system.

If the local Unified CCE or Unified CCX child system were to fail, the Unified CCE gateway PG would not be able to connect to it, and the Unified ICM parent would then consider all of the agents to be off-line and not available. If calls were sent to the local Unified CM while the child Unified CCE or Unified CCX system was down, the call-forward-on-failure processing would take over the call for the CTI route point. This method would redirect the call to another site or an answering resource to play a message telling the caller there was an error and to call again later.

Parent/Child Reporting and Configuration Impacts

During any time that the Unified CCE child is disconnected from the Unified ICM parent, the local IP-ACD is still collecting reporting data and allows local users to make changes to the child routing scripts and configuration. The Unified CCE gateway PG at the child site will cache these objects and store them in memory (and eventually to disk) to be sent later to the Unified ICM parent when it is available. This functionality is available only if the Unified CCE gateway PG is co-located at the child Unified CCE site.

Other Considerations for the Parent/Child Model

Multi-channel components such as EIM/WIM and Unified Outbound Option may be installed only at the child Unified CCE level, not at the parent. They are treated as nodal implementations on a site-by-site basis.

Other Considerations for High Availability

A Unified CCE failover can affect other parts of the solution. Although Unified CCE may stay up and running, some data could be lost during its failover, or other products that depend on Unified CCE to function properly might not be able to handle a Unified CCE failover. This section examines what happens to other critical areas in the Unified CCE solution during and after failover.

Reporting

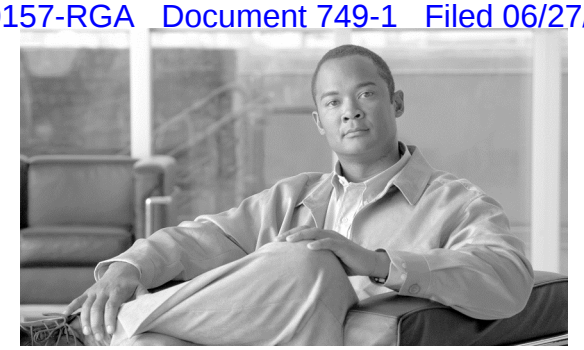
The Unified CCE reporting feature uses real-time, five-minute and reporting interval (15 or 30 minute) data to build its reporting database. Therefore, at the end of each five-minute and reporting interval (15 or 30 minute), each Peripheral Gateway will gather the data it has kept locally and send it to the Call Routers. The Call Routers process the data and send it to their local Logger for historical data storage. That data is then replicated to the HDS database from the Logger as it is written to the Logger database.

The Peripheral Gateways provide buffering (in memory and on disk) of the five-minute and reporting interval (15 or 30 minute) data collected by the system to handle network connectivity failures or slow network response as well as automatic retransmission of data when the network service is restored. However, physical failure of both Peripheral Gateways in a redundant pair can result in loss of the half-hour or five-minute data that has not been transmitted to the Central Controller. Use redundant Peripheral Gateways to reduce the chance of losing both physical hardware devices and their associated data during an outage window.

When agents log out, all their reporting statistics stop. The next time the agents log in, their real-time statistics start from zero. Typically, Central Controller failover does not force the agents to log out or reset their statistics; however, if the PG fails-over, their agent statistics are reset because the PIM and OPC processes that maintain these values in memory are restarted. If the CTI OS or CAD servers do not fail-over or restart, the agent desktop functionality is restored to its pre-failover state.

For further information, refer to the *Reporting Guide for Cisco IPCC Enterprise & Hosted Editions*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps4145/products_user_guide_list.html



CHAPTER 4

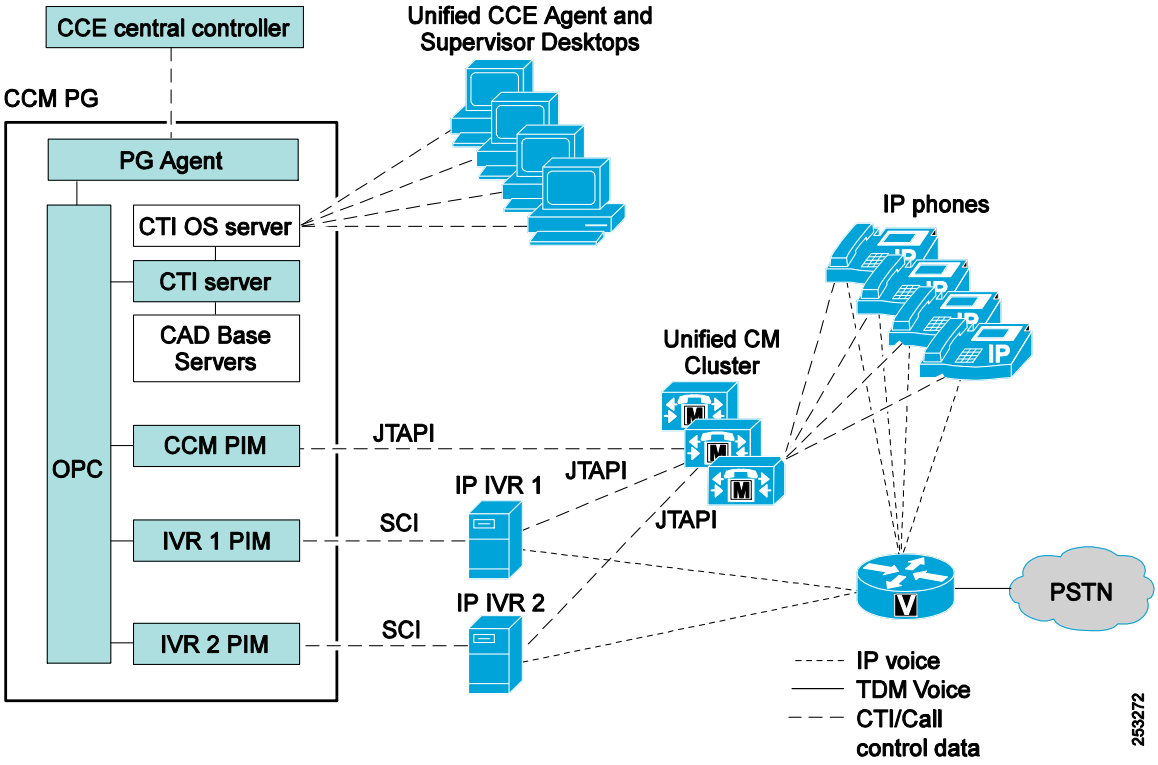
Unified Contact Center Enterprise Desktop

The Cisco Unified Contact Center Enterprise (CCE) solution delivers a comprehensive set of desktop applications and services.

Desktop Components

The desktop applications themselves typically run on Agent desktops, Supervisor desktops, Administration & Data Servers or Administration Client. Services supporting the desktop applications typically run on the Unified CCE Peripheral Gateway (PG) server. Within the Unified CCE deployment, there may be one or more PG systems, and for each PG there is one set of active desktop services, which includes the CTI Object Server (CTI OS) and the Cisco Agent Desktop Base Services for Cisco Agent Desktop (CAD) deployments. Figure 4-1 depicts the components within a Unified CCE deployment that support the various desktop applications.

Figure 4-1 Generic Unified CCE Desktop Components



In the Unified CCE solution, the Peripheral Gateway may be deployed in either a simplex or duplex configuration. (Simplex mode is not supported for production environments.) Duplex configurations provide redundant desktop services for failover recovery support. These systems are typically identified as the primary, or A-side, and the backup, or B-side. For production deployments, a duplex configuration is required.

CTI Object Server

The CTI Object Server (CTI OS) is a high-performance, scalable, fault-tolerant, server-based solution for deploying CTI applications. CTI OS is a required component for CTI Toolkit desktop and Cisco Agent Desktop (CAD) solutions and is Cisco's latest version of the CTI implementation.

Communications from the desktop applications, such as agent state change requests and call control, are passed to the CTI OS server running on the Cisco Unified Peripheral Gateway. CTI OS serves as a single point of integration for CAD desktops, CTI Toolkit desktops, and third-party applications such as Customer Relationship Management (CRM) systems, data mining, and workflow solutions.

The CTI Object Server connects to CTI Server via TCP/IP and forwards call control and agent requests to CTI Server, which in turn forwards to the Open Peripheral Controller (OPC). From there, depending on the type of request, OPC will forward to either the CCM Peripheral Interface Manager (PIM) or to the Unified CCE Central Controller.

Requests initiated from the desktop application that affect the agent state are sent to the Unified CCE Central Controller, while requests initiated from the desktop application that affect call control are sent to the CCM PIM. The Unified CCE Central Controller monitors the agent state so that it knows when it can and cannot route calls to that agent and can report on that agent's activities.

Call control flows from the agent desktop application to Cisco Unified Communications Manager (Unified CM). Unified CM then performs the requested call or device control. The desktop services located on the PG keep the agent desktop application synchronized with the agent's IP phone state.

CTI Toolkit desktop configuration and behavior information is also managed at the CTI OS server, simplifying customization, updates, and maintenance, and supporting remote management.

CTI Object Server Services:

- Desktop Security — Supports secure socket connections between the CTI Object Server on the PG and the agent, supervisor, or administrator desktop PC. Any CTI application built using the CTI Toolkit C++ Client Interface Library (CIL) Software Development Kit (SDK) can utilize the desktop security feature.

Note Desktop Security is not currently available in the .NET and Java CILs.

- Quality of Service (QoS) — Supports packet prioritization with the network for desktop call control messages.

Note QoS is not currently available in the .NET and Java CILs.

- Failover Recovery — Supports automatic agent login upon failover.
- Chat — Supports message passing and the text chat feature between agents and supervisors.
- Silent Monitoring — Supports VoIP monitoring of active calls. The CTI Object Server communicates with the Silent Monitor Service (SMS) to start/stop the VoIP packet stream forwarding.

The CTI Object Server is typically installed in duplex mode, with two CTI OS servers running in parallel for redundancy, one on PG side-A and one on PG side-B. The CTI Toolkit Desktop applications randomly connect to either server and automatically fail-over to the alternate server if the connection to the original CTI OS server fails. CTI OS can also run in simplex mode with all clients connecting to a single server, but Cisco does not recommend this configuration. (Simplex mode is not supported for production environments.)

Agent capacity sizing for the PG is covered in the chapter Sizing Unified CCE Components and Servers, page 10-1.

Note The CTI OS server interfaces to any desktop application built using the CTI Desktop Toolkit Software Development Kit. A single CTI OS server can support the use of both CAD and CTI Toolkit desktops concurrently.

CAD Base Services

Cisco Agent Desktop (CAD) is a software suite that provides a feature-rich packaged solution. CAD consists of user applications and the CAD Base Services, which can run co-resident on the Peripheral Gateway within a Unified CCE deployment and are required for CAD deployments only. The CAD Base Services provide redundancy and warm standby capabilities.

CAD Base Services:

- Cisco Chat Service — Supports message passing and the text chat feature.

- Cisco Enterprise Service — Communicates with the Unified CCE components to provide call data to the user applications.
- Cisco Browser and IP Phone Agent Service — Provides services for CAD-BE and IPPA agent applications.
- Cisco Synchronization Service — Synchronizes the Unified CCE and CAD-specific configuration data.
- Cisco LDAP Monitor Service — Manages the storage and retrieval of CAD configuration data.
- Cisco Recording and Statistics Service — Manages the storage and retrieval of call recording, agent call, and agent state change data used in reports.
- Cisco Licensing and Resource Manager Service — Manages user licenses and controls failover behavior.
- Cisco Recording and Playback Service — Provides the call recording and playback feature.
- Cisco VoIP Monitor Service — Provides the voice streams for the call recording and silent monitoring features if server-based monitoring is used.

For more information on CAD, refer to the product documentation available at

http://www.cisco.com/en/US/products/sw/custcosw/ps427/tsd_products_support_series_home.html

Cisco Unified Contact Center Enterprise (Unified CCE) supports a variety of desktop application choices for agents and supervisors, as described in the following sections.

Agent Desktops

An agent desktop application is a required component of a Unified CCE deployment. The contact center agent uses this application to perform agent state control (login, logout, ready, not ready, and wrap-up) and call control (answer, release, hold, retrieve, make call, transfer, and conference). In addition to these required features, the application can provide enhanced features that are useful in a contact center environment.

Cisco offers the following primary types of Unified CCE agent desktop applications:

- Cisco Agent Desktop (CAD) — A packaged agent desktop solution supporting an embedded browser and scripted workflow automation.
- CTI Desktop Toolkit — A development toolkit that provides agent desktop applications and that supports full customization and integration with other applications, customer databases, and Customer Relationship Management (CRM) applications.
- Cisco Unified CRM Connector for Siebel — A CTI driver for the Siebel Communication Server.
- Cisco Unified IP Phone Agent — An agent desktop solution provided through the Cisco Unified IP Phone display.
- Cisco Agent Desktop Browser Edition (CAD-BE) — A browser-based agent application that supports many of the features of the CAD windows-based agent application with lower platform requirements.
- Agent Desktop Applications Offered through Cisco partners:
- Partner Agent Desktops — Custom agent desktop applications are available through Cisco Technology Partners. These applications are based on the CTI Desktop Toolkit and are not discussed individually in this document.

- Prepackaged CRM integrations — CRM integrations are available through Cisco Unified CRM Technology Partners. They are based on the CTI Desktop Toolkit and are not discussed individually in this document.

Agent Mobility

Within the Unified CCE deployment, the agent desktop application is not statically associated with any specific agent or IP phone extension. Agents and phone extensions (device targets) are configured within the Unified CCE configuration and associated with a specific Unified CM cluster.

When logging in from an agent desktop application, the agent is presented with a dialog box that prompts for agent ID or login name, password, and the phone extension to be used for that session. At that time the agent ID, phone extension, and agent desktop IP address are dynamically associated. The association is released when the agent logs out.

This mechanism enables an agent to work (or *hot-desk*) at any workstation. It also enables agents to take their laptops to any Cisco Unified IP Phone and log in from that device (assuming the phone has been configured in Unified CCE and in Unified CM to be used in the Unified CCE deployment). Agents can also log in to other phones using the Cisco Extension Mobility feature. For more information on Extension Mobility, refer to the Extension Mobility section of the *Cisco Unified Communications Manager Features and Services Guide*, available at

http://www.cisco.com/en/US/products/sw/voicesw/ps556/prod_maintenance_guides_list.html

Supervisor Desktops

In addition to the agent desktop application, a supervisor desktop application is also available. The contact center supervisor uses this application to monitor agent state for members within their team. The supervisor desktop also enables Silent Monitoring of agents during active calls.

Cisco offers the following types of Unified CCE supervisor desktop applications:.

- Cisco Supervisor Desktop (CSD) — A packaged supervisor desktop solution.
- CTI Desktop Toolkit — A development toolkit that provides a supervisor desktop application and supports customization and integration with other applications, customer databases, and Customer Relationship Management (CRM) applications.
- Supervisor Desktop Applications Offered through Cisco partners
- Prepackaged CRM integrations — CRM integrations are available through Cisco Unified CRM Technology Partners. They are based on the CTI Desktop Toolkit and are not discussed individually in this document.

Desktop Solutions

Depending on the requirements of the contact center, a particular type of desktop might be better suited to the solution. Table 4-2 contains an abbreviated list of the functionality available in the various desktop applications. It is intended to provide a starting point to determine the desktop that best meets specific solution requirements. Further information is available for each of the Cisco desktops in the sections below and in their respective product specifications at http://www.cisco.com.

Table 4-1 Features Supported by Cisco Desktop Solutions

Desktop Functionality	Cisco Agent Desktop	Cisco Agent Desktop Browser Edition	CTI Desktop Toolkit	Cisco Unified CRM Connector for Siebel	IP Phone Agent
Turn-kep desktop applications	Yes	Yes	Yes	Yes	Yes
Custom desktop development using C++, .NET, and Java			Yes		
Desktop Security	Yes		Yes		
Workflow Automation	Yes	Yes			
Mobile (Remote) Agents	Yes	Yes	Yes		
Siebel Integration				Yes	
Silent Monitoring	Yes	Yes	Yes		Yes
Integrated Recording Capacity	Yes	Yes			Yes
Monitor Mode Applications			Yes		
Outbound Calls	Yes		Yes		
Microsoft Terminal Services Support	Yes		Yes		
Citrix Presentation Server Support	Yes		Yes		
Agent Mobility	Yes	Yes	Yes		Yes
IP Phone Solution (no soft desktop)					Yes
Specific capability or integration not offered by Cisco					

Cisco Agent Desktop Solution

The Cisco Agent Desktop (CAD) solution is a suite of packaged desktop applications and services. CAD offers a rich set of features for the contact center environment, including:

- Agent state and call control

  Agent Desktop provides call control capabilities (call answer, hold, conference, and transfer) and ACD state control (ready/not ready, wrap up, and so forth).

- Work flow automation

  The work flow automation feature allows an administrator to customize the agent environment and how the user applications interact with that environment. Work flow automation enables data processing actions to be scheduled based on telephony events (for example, popping data into a third-party application on the answer event and sending email on the dropped event). Work flow automation interfaces with applications written for Microsoft Windows browsers and terminal emulators. Some customizations can be as simple as using keystroke macros for screen pops.

- On-demand recording

  The supervisor (and, if enabled, the agent) can record a customer phone call for later review by a supervisor.

- Cisco IP Phone Agent service

  With this XML service, agents using Cisco IP phones can log in and use their phone to perform most of the agent functions found in an agent desktop application.

- Collaboration

  A supervisor can text-chat directly with agents or agent teams. Agents can text-chat with supervisors or other team members (if enabled). The supervisor can push web pages to agents and send team messages to agent desktops. This interactive collaboration enables the contact center to communicate better, increase productivity, improve customer responsiveness, and coach or train agents.

- Task automation

  Routine agent tasks, such as email, conferences to knowledge workers, launching other applications, and high-priority chat, can be configured as task buttons on the agent's toolbar to reduce call duration and improve customer responsiveness.

- Silent monitoring

  Supervisors can initiate a silent monitoring session with an agent on their team.

What's New In This Version

CAD 8.0 includes these new features:

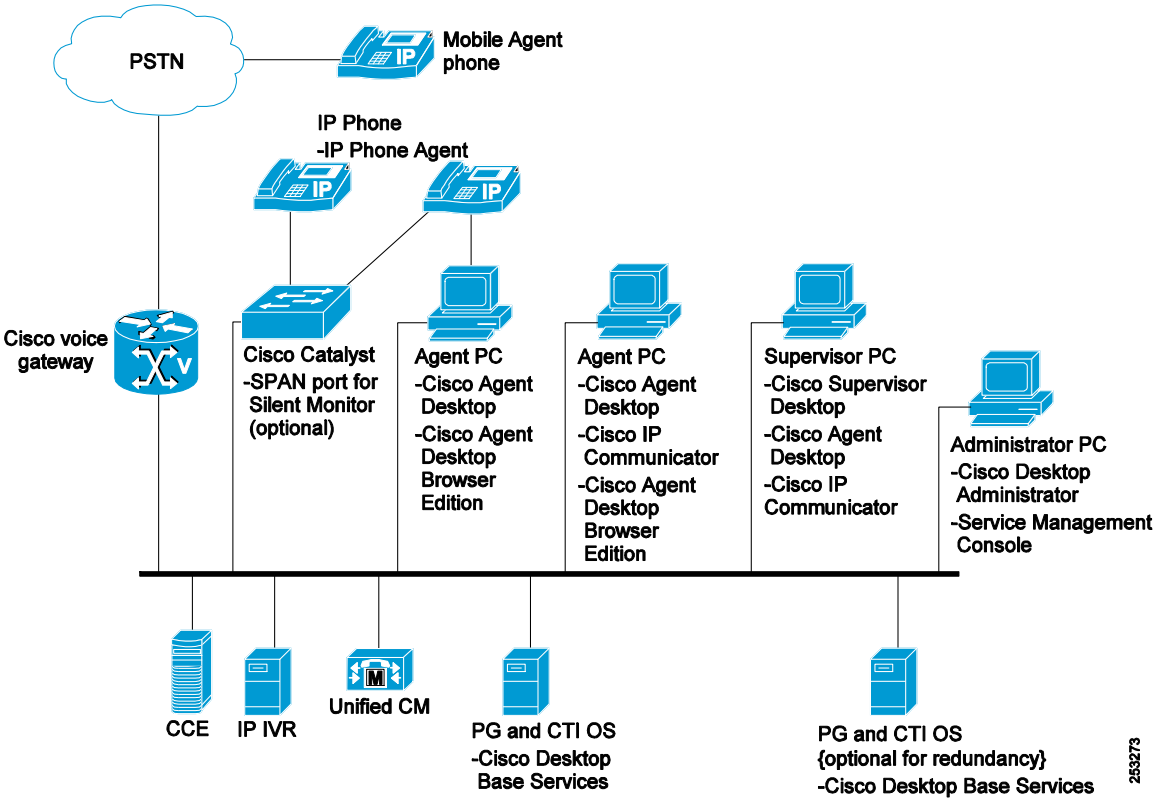
- Support for agent phones enabled with multi-line extensions
- Support for extension mobility cross-cluster (EMCC)
- Support for multi-byte languages in Cisco IP Phone Agent service
- Portions of Cisco Desktop Administrator functionality moved to a web-based application
- CAD-BE support for Mac OS
- Independent client configuration packaging
- Polish and Turkish localizations
- CAD co-resident on a PG running VMware
- CAD support for a third-party recording action
- Replacement of MSDE with flat files

CAD User Applications

CAD user applications include the following applications for contact center agents, supervisors, and administrators. (See Figure 4-2.)

- Cisco Agent Desktop: Windows-based agent application
- Cisco Agent Desktop–Browser Edition (CAD-BE): Java-based version of Agent Desktop
- Cisco IP Phone Agent (IPPA): IP phone service agent application
- Cisco Supervisor Desktop (CSD): Windows-based supervisor application
- Cisco Desktop Administrator (CDA): Web-based administrative application
- Cisco Desktop Work Flow Administrator: Windows-based work flow configuration tool

Figure 4-2 Cisco Agent Desktop System Configuration and Components



CAD Application Features

Table 4-2 compares some of the more important CAD features to assist users in selecting the appropriate agent application for their deployment.

Table 4-2 Comparison of Major CAD Features

Feature	CAD	CAD-BE	IPPA
Call control	Yes	Yes	n/a**
VPN/Remote Agent support	Yes	Yes	Yes
Chat / Unified Presence Integration	Yes	No	No
Supports Cisco IP Communicator	Yes	Yes	No
Team Messages	Yes	No	No
Supports Mobile Agent	Yes	Yes	n/a
Real Time Queue and Agent Displays	Yes	No	Yes
Supports Cisco Outbound Dialer	Yes	No	No
Integrated browser	Yes	Yes	n/a
Call event work flow automation	Yes	Yes (limited)	No
Agent state work flow automation	Yes	No	No
Supports thin client environment	Yes	No	n/a

Table 4-2 Comparison of Major CAD Features (continued)

Desktop monitoring and recording*	Yes	No	No
SPAN monitoring and recording*	Yes	Yes	No
Unified CM monitoring and recording*	Yes	Yes	Yes

*** For more detailed information on supported recording and monitoring, refer to *Configuring and Troubleshooting VoIP Monitoring*, available at http://www.cisco.com/en/US/products/sw/custcosw/ps427/prod_troubleshooting_guides_list.html.**

** Call control actions are performed by using the IP phone's call control softkeys.

For more information on CAD agent applications, refer to the appropriate user guide, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps427/tsd_products_support_series_home.html

Cisco Agent Desktop

Cisco Agent Desktop is a Windows application that runs on the agent PC. It works with either a hardware IP phone or the Cisco IP Communicator soft phone. Agent Desktop interfaces with the CTI OS service for call control and agent state change events; for all other features, it communicates with the CAD services.

Agent Desktop supports Desktop, SPAN, and Unified CM monitoring and recording.

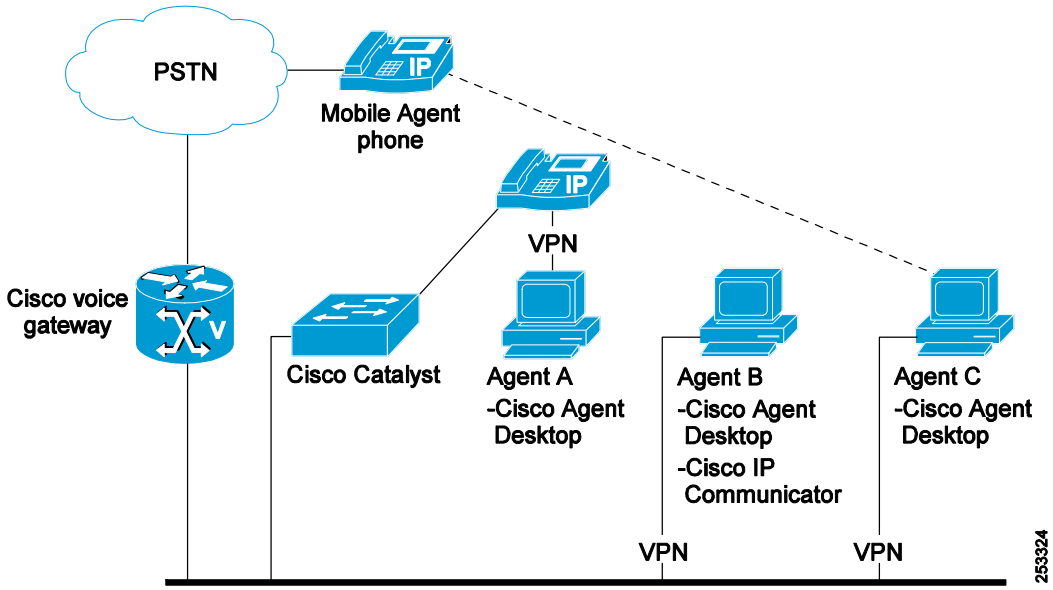
Figure 4-3 illustrates various ways agent desktops can be configured in a contact center.

- Agent A shows an agent who uses a hardware IP phone. The IP phone connects directly to the agent's PC via a network cable. This is the configuration required for desktop monitoring. CAD supports a VPN connection between the agent's PC and the contact center network.
- Agent B shows an agent who uses Cisco IP Communicator. This configuration also supports a VPN connection to the contact center network. This is the most common configuration for remote agents.
- Agent C shows Agent Desktop used with the Mobile Agent feature. Mobile agents are agents whose phones are not directly controlled by Unified CM. Agents might use their home phones or cell phones as their agent device. In this case, the agent provides a CTI port to associate with their remote phone when logging in. ACD calls for the logged-in agent are sent to the CTI port, which causes the call to appear at the mobile agent's phone device. There is a logical relationship (the dashed line) between the agent and the mobile phone. CAD supports a VPN connection between the agent and the contact center network in this configuration. Mobile agents can be monitored and recorded using SPAN monitoring.

For more information about Cisco Agent Desktop features and capabilities, refer to the *Cisco Agent Desktop User Guide,* available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html

Figure 4-3 CAD Agents and Components



Cisco Agent Desktop Browser Edition

Cisco Agent Desktop—Browser Edition (CAD-BE) is a Java applet that runs in Microsoft Internet Explorer (Windows machines) or in Mozilla Firefox (Windows and Linux machines). CAD-BE interfaces with the BIPPA service for call control; BIPPA in turn interfaces with the CTI service. For all other features it communicates with the CAD services.

CAD-BE can use the abilities and features of any Java applet. This enables CAD-BE to have a small footprint on the agent desktop and still provide essential agent features.

Some limitations of CAD-BE beyond those listed in Table 4-2 include:

- Desktop monitoring and recording is not supported. CAD-BE does support server monitoring and recording and Unified CM monitoring.
- The integrated browser is external to the CAD-BE window and is launched when CAD-BE is launched.
- The “click to dial from browser” feature is not supported.
- Agents cannot modify enterprise data.
- The dial pad does not support dial string formatting, phone books, or recent call list.
- Window behavior (for example, stealth or always on top) cannot be configured by the agent.
- For more information about CAD-BE features and capabilities, refer to the *Cisco Agent Desktop—Browser Edition User Guide,* available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html.

Cisco Unified IP Phone Agent

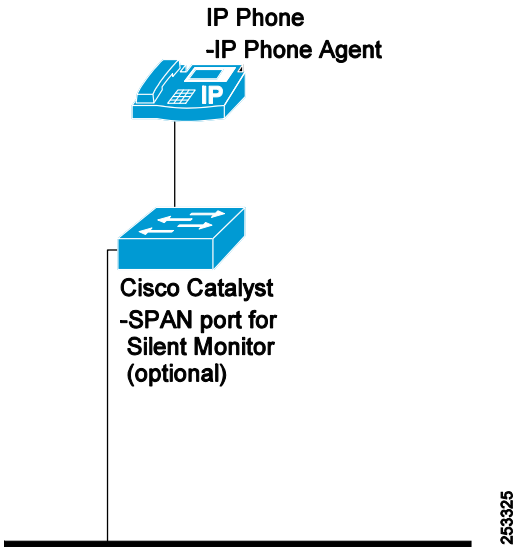
Cisco IP Phone Agent (IPPA) runs as an IP phone XML service. The agent is not required to have a PC. IPPA includes all the basic features required by a contact center agent, as well as advanced features such as reason codes, wrap-up data, and on-demand recording.

IPPA agents can be monitored and recorded using server monitoring, and monitored using Unified CM monitoring.

For more information about IPPA features and capabilities, refer to the *Cisco IP Phone Agent User Guide* available at http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html.

The following figure ilustrates the components used by IP Phone agents.

Figure 4-4 Cisco IP Phone Agent Components



Cisco Supervisor Desktop

Cisco Supervisor Desktop provides a graphical view of the agent teams managed by the supervisor. An expandable navigation tree, similar to that in Windows Explorer, is used to navigate to and manage team resources.

Supervisors are able to view real-time information about the agents in a team as well as interact with those agents. The supervisor can:

- View and change an agent's state
- View contact information specific to the agent
- Silently monitor and/or record the agent's calls
- Barge-in or intercept an agent's call
- Chat with the agent using an instant message window
- Push a web page to the agent's desktop

When Supervisor Desktop is installed, an instance of Agent Desktop is installed as well. Agent Desktop is needed by the supervisor in order to take calls, barge in, intercept, and retrieve skill group statistics.

The Supervisor Work Flow module enables configurable actions to be triggered when specific events occur in the contact center. For example, a supervisor work flow can be set up so that whenever more than ten calls are in queue for a specified skill group, an audible alert sounds and the skill group name is highlighted in red on the supervisor's desktop. Another work flow sends an email to specified email addresses when certain events occur. The email contains information related to the condition that caused the event, as well as custom text.

Supervisors can use the Supervisor Record Viewer to review recordings and mark selected recordings for extended retention. The supervisor can also save recordings for permanent retention in a format that can be played by any media player.

For more information about Supervisor Desktop features and capabilities, refer to the *Cisco Supervisor Desktop User Guide*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html.

Cisco Desktop Administrator

Cisco Desktop Administrator enables an administrator to configure the CAD services and CAD client applications. Individual work flow groups containing agents and supervisors can be configured separately to provide specific functionality to particular groups of agents.

Desktop Administrator consists of two components:

- Cisco Desktop Work Flow Administrator, a Windows-based application
- Cisco Desktop Administrator, a web-based application

Cisco Desktop Work Flow Administrator is used to configure the following:

- Dial strings
- Phone books
- Reason codes
- Wrap-up data
- Record/monitor notification
- Work flow groups

Dial strings, phone books, reason codes, and wrap-up data can be configured on the global and work flow group level.

Work flows and user interfaces can be configured for specific agent types (CAD agents and CAD-BE agents).

Cisco Desktop Administrator is used to configure the following:

- Enterprise data fields and layouts
- Silent monitoring and recording
- Personnel and assigning users to work flow groups
- Cisco Unified Presence settings

For more information about Cisco Desktop Administrator features and capabilities, refer to the *Cisco Desktop Administrator User Guide,* available at

- http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_user_guide_list.html.

Cisco Desktop Monitoring Console

The Cisco Desktop Monitoring Console is a Java application that monitors the status of the CAD services. It provides a convenient interface for an administrator to use to get real-time information about the CAD system.

CTI Desktop Toolkit Solution

The CTI Desktop Toolkit provides a Software Development Kit (SDK) for custom development of desktop applications. The CTI Desktop Toolkit supports C++, Java, and .NET development Client Interface Libraries (CILs) and provides sample applications for customization.

Additionally, the CTI Desktop Toolkit ships complete with pre-built, ready-to-run agent desktop, supervisor desktop, and call center monitoring applications. These applications can be used as-is or can be customized further to meet the particular needs of a call center.

The CTI Desktop Toolkit also offers advanced tools for integrating desktop applications with a database, Customer Relation Management (CRM) applications, or other contact center applications.

The CTI Toolkit Desktop solution offers a rich set of features for the contact center environment, including:

- Collaboration — A supervisor can text-chat directly with agents, and agents can text-chat with supervisors or other team members (if enabled). Interactive collaboration enables the contact center to communicate better, increase productivity, improve customer responsiveness, and coach or train agents.
- Secure Desktop Connection — Desktop security is provided between the agent desktop and the CTI OS server.
- Silent Monitoring — A supervisor can initiate a silent monitoring session with an agent within their team.

CTI Toolkit Software Development Kits and User Applications

The CTI Desktop Toolkit provides the following user tools and applications. The CTI Desktop Tookit Monitor mode applications (i.e AllAgents) can be used only if the number of Skill per Agent is less than 20.

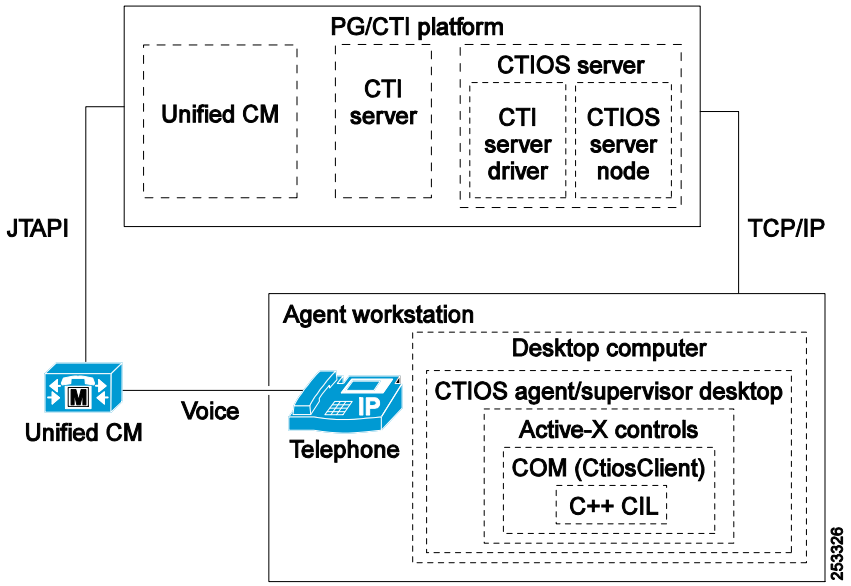
- C++ CIL API — A Windows software development kit for developing C++ CTI applications
- Java CIL API — A cross-platform library for developing Java CTI applications
- .NET CIL API — A Windows software development kit for developing custom .NET framework CTI applications
- COM CIL API — A set of COM Dynamic Link Libraries (COM DLL) for building a Visual Basic 6.0 CTI application
- ActiveX Controls — A set of Windows GUI controls for custom desktop development using Microsoft Visual Basic 6.0
- CTI OS Runtime Callable Wrappers — A set of .NET assemblies that allows the use of COM CIL and ActiveX controls in native .NET applications
- CTI Toolkit Agent Desktop — A Windows Visual Basic application built upon the COM CIL and Active-X controls, providing agent desktop functionality
- CTI Toolkit Supervisor Desktop — A Windows Visual Basic application built upon the COM CIL and Active-X controls, providing supervisor desktop functionality
- CTI Toolkit Outbound Desktop — A Windows Visual Basic application built upon the COM CIL and Active-X controls, supporting outbound call center campaigns in addition to standard agent desktop functionality
- CTI Toolkit Combo Desktop — A Windows agent and supervisor application based on the .NET CIL, which combines support for agent, supervisor, and outbound functionality

- CTI Toolkit All-Agents Monitor — A Windows Admin application based on the C++ CIL, providing call center agent status monitoring
- CTI Toolkit All-Calls Monitor — A Windows Admin application based on the C++ CIL, providing call center call status monitoring

Figure 4-5 illustrates the architecture of the CTI Desktop Toolkit. For more information regarding the CTI Desktop Toolkit, refer to the *CTI OS Developer's Guide for Cisco ICM/IPCC Enterprise and Hosted Editions*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_programming_reference_guides_list.html

Figure 4-5 CTI Desktop Toolkit Architecture



C++ CIL API

The CTI Desktop Toolkit C++ CIL provides a set of header files and static libraries for building C++ CTI applications using Microsoft Visual Studio .NET. The C++ CIL also supports a secure desktop connection between the agent PC and the CTI Object Server on the PG.

Java CIL API

The CTI Desktop Toolkit Java CIL provides a powerful cross-platform library for developing Java CTI applications.

.NET CIL API

The CTI Desktop Toolkit .NET CIL provides native .NET class libraries for developing native .NET Framework applications. The .NET Combo Desktop is provided as a sample application built using the .NET CIL.

COM CIL API

The CTI Desktop Toolkit COM CIL provides a set of COM Dynamic Link Libraries for building Visual Basic 6.0 CTI applications. The CTI Toolkit Agent and Supervisor Desktops are provided as sample applications built with Visual Basic 6.0 and using the COM CIL.

ActiveX Controls

The CTI Toolkit includes a set of ActiveX controls to enable rapid application development. The ActiveX controls are UI components that enable easy drag-and-drop creation of custom CTI applications in a variety of container applications. Container applications include Microsoft Visual Basic 6.0, Microsoft Internet Explorer, Microsoft Visual C++ 7.0, Borland Delphi, Sybase Powerbuilder, and other applications supporting the OC96 ActiveX standard.

The ActiveX Controls include:

- Agent State Control
- Chat Control
- Emergency Assist Control
- Alternate Control
- Answer Control
- Bad Line Control
- Call Appearance Control
- Conference Control
- Hold Control
- Make Call Control
- Reconnect Control
- Status Bar Control
- Record Control
- Transfer Control
- Agent Statistics Control
- Skill Group Statistics Control
- Agent Select Control
- Supervisor Control
- Silent Monitor Control

CTI Toolkit Agent Desktop

The CTI Toolkit Agent Desktop is a Microsoft Windows application that runs on an agent's desktop PC and works with either a hardware IP phone or the Cisco IP Communicator software phone. The CTI Toolkit Agent Desktop interfaces with the CTI OS server for call control and agent state change events.

The CTI Toolkit Agent Desktop includes support for desktop monitoring, which captures the voice stream on the agent's IP phone to support the silent monitoring and call recording features.

CTI Toolkit Supervisor Desktop

The CTI Toolkit Supervisor Desktop is a Microsoft Windows application that runs on a supervisor's desktop PC. The CTI Toolkit Supervisor Desktop interfaces with the CTI OS server for agent state change events and real-time statistics updates. The CTI Toolkit Supervisor Desktop provides the contact center supervisor with the ability to manage a team of agents. Supervisors are able to view real-time information about the agents in a team as well as interact with these agents. A supervisor can select an agent to change the agent's state, view information specific to that agent, silently monitor the agent's call, barge in or intercept the agent's call, or chat with the agent.

A supervisor may also receive emergency assistance requests from agents on their team through the supervisor desktop.

In Unified CCE, supervisors may also be configured to act as agents. When this is done, the standard set of agent phone controls is available on the Supervisor Desktop.

CTI Toolkit Outbound Desktop

The CTI Toolkit Outbound Desktop is a Microsoft Windows application that runs on an agent's desktop PC and works with either a hardware IP phone or the Cisco IP Communicator software phone. The CTI Toolkit Outbound Desktop interfaces with the CTI OS server for call control and agent state change events. In addition to the standard set of agent controls present in the CTI Toolkit Agent Desktop, the Outbound Desktop provides a set of controls for managing outbound call campaigns. Outbound calls are automatically managed by Unified CCE, and the agent utilizes the additional controls to accept the next outbound call.

CTI Toolkit Combo Desktop

The CTI Toolkit Combo Desktop is a Microsoft Windows .NET application that runs on an agent's desktop PC and works with either a hardware IP phone or the Cisco IP Communicator software phone. The CTI Toolkit Combo Desktop interfaces with the CTI OS server for call control and agent state change events.

The Combo Desktop integrates the functionality of the Toolkit Agent, Supervisor, and Outbound desktops into a single .NET application. The Combo Desktop source code is also provided as a starting point for custom desktop development using the Microsoft .NET Framework.

CTI Toolkit All-Agents Monitor

The CTI Desktop Toolkit ships complete with a ready-to-run All-Agents Monitor application. This application provides a call center administrator with the ability to monitor agent login and state activity within the call center.

CTI Toolkit All-Calls Monitor

The CTI Desktop Toolkit ships complete with a ready-to-run All-Calls Monitor application. This application provides a call center administrator with the ability to monitor call activity within the call center.

Cisco Unified CRM Connector for Siebel Solution

The Cisco Unified CRM Connector for Siebel is an installable component developed by Cisco that enables integration of the Cisco Unified CCE with the Siebel CRM Environment. In this solution, the Siebel Agent Desktop provides the agent state and call control interface. The Siebel Desktop utilizes the Cisco Unified CRM Connector for Siebel, which is built on top of the CTI Desktop Toolkit C++ CIL to communicate with the CTI Object Server.

For more information on the capability of the Siebel eBusiness solution, refer to the Siebel website at

http://www.siebel.com/index.shtm

Deployment Considerations

This section covers the deployment considerations.

Citrix and Microsoft Terminal Services (MTS)

This section discusses deploying Cisco Agent Desktop and Cisco Toolkit Desktop in a Citrix or Microsoft Terminal Services (MTS) environment.

Cisco Agent Desktop

Cisco Unified CCE supports running Cisco Agent Desktop within a Citrix terminal services environment. When planning to use Citrix terminal services for CAD, take the following considerations into account:

- Cisco Supervisor Desktop (CSD) and Cisco Desktop Administrator (CDA) are not supported in a Citrix terminal services environment.
- Desktop monitoring (for silent monitoring and recording) is not supported with Citrix terminal services. SPAN port monitoring must be used instead.
- Macros work only if they involve applications running on the Citrix server, and not those running on the client PC.
- Only one Citrix user name is supported per CAD application login.
- The login ID and extension that appear by default in the login dialog box when CAD is started, are those associated with the last login by any user.
- The Citrix web client is not supported.
- Only Citrix 4.0 and 4.5 running on Windows 2000 Server or Windows 2003 Server are supported.

For implementation details, refer to *Integrating CAD into a Citrix MetaFrame Presentation Server or Microsoft Terminal Services Environment*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_implementation_design_guides_list.html

Cisco Toolkit Desktop

Cisco Unified CCE supports running CTI Toolkit Desktop within the Citrix and Microsoft Terminal (MTS) Services environments. When planning to use Citrix terminal services with the CTI Toolkit Desktop, take into account the following considerations:

- Versions of Citrix MetaFrame Presentation Server prior to Version 4.0 or 4.5 are not supported. Earlier versions have limitations for publishing Microsoft .NET applications.
- CTI OS Java CIL client applications are supported only on Citrix MetaFrame Presentation Server 4.0 and 4.5 for the Windows platform. There is no planned support for Citrix MetaFrame Presentation Server 4.0 or 4.5 on UNIX.
- Silent Monitoring is supported within a Citrix or MTS environment.
- CTI OS Client Desktop sounds such as dial tones and DTMF tones are not audible.

For implementation details, refer to *Integrating CAD into a Citrix MetaFrame Presentation Server or Microsoft Terminal Services Environment*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_implementation_design_guides_list.html

Silent Monitoring

Silent monitoring enables supervisors to monitor the conversations of agents within their team. Supervisors are not able to participate actively in the conversation, and the agent(s) and caller(s) are unaware they are being monitored. Both the Cisco Agent Desktop and the CTI Desktop Toolkit provide solutions support for silent monitoring. CAD Server-based monitoring supports Agent Desktops, IP Phone Agents, and Mobile Agents. Desktop monitoring supports only desktop agents. CTI OS releases 7.2 and later support two types of silent monitors: CTI OS silent monitor and Unified CM silent monitor.

CTI OS silent monitoring is accomplished via one or more VoIP monitoring services located either on the agent's desktop (desktop monitoring) or on a separate VoIP monitor server (server-based monitoring). CTI OS uses server-based silent monitoring to support mobile agents and desktop-based silent monitoring to support traditional (non-mobile) Unified CCE agents.

Unified CM accomplishes silent monitoring with a call between the supervisor's (monitoring) device and agent's (monitored) device. The agent's phone mixes and sends the agent's conversation to the supervisor's phone, where it is played out to the supervisor. Unified CM silent monitoring can be initiated by any of the CTI OS supervisor desktops (out-of-the-box, Java, or .NET). Any Unified CCE agent desktop, including Siebel, can be silently monitored using Unified CM silent monitoring, provided the following requirements are met:

- The agent to be silently monitored is using a Cisco Unified IP Phone 7941, 7942, 7945, 7961, 7962, 7965, 7970, 7971, or 7975, and Cisco IP Communicator 7.0 or later.
- The contact center is using Cisco Unified CM 6.0 or higher (Support for IP communicator requires Cisco Unified CM 6.1(3) or higher).
- Phones are configured to use RTP streams (SRTP streams cannot be silently monitored).

Unified CM silent monitoring does not support mobile agents.

Unified CM silent monitoring supports a maximum of one silent monitoring session and one recording session for the same agent phone.

Supervisors can use any Cisco IP Phone, including Cisco IP Communicator, to silently monitor.

Note G.722 is used as the default codec for regions that are configured for G.711 on devices that support G.722. However, G.722 is not supported with silent monitoring and call recording based on CAD, CTI OS, or Unified CM. To disable this default, in Unified CM Administration go to **Enterprise Parameters** and set **Advertise G.722 Codec** to **disabled**.

CTI Toolkit Silent Monitor

A given CTI OS Server can be configured to use either CTI OS silent monitor or Unified CM silent monitor, or to disable silent monitoring. When supervisor desktops connect to the CTI OS Server, this configuration is downloaded. The supervisor desktop uses this information to invoke the configured type of silent monitor when the Start Silent Monitor button is pressed. The initial message from the supervisor desktop is used by the CTI OS Server to drive either the CTI OS or Unified CM silent monitor.

For details regarding the configuration of silent monitoring, system administrators can refer to the *CTI OS System Manager's Guide for Cisco ICM/IPCC Enterprise & Hosted Editions*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/prod_installation_guides_list.html

Developers implementing either the CTI OS or Unified CM silent monitor should refer to the *CTI OS Developer's Guide for Cisco ICM/IPCC Enterprise & Hosted Editions*, available at

http://www.cisco.com/en/US/products/sw/custcosw/ps1844/products_programming_reference_guides_list.html

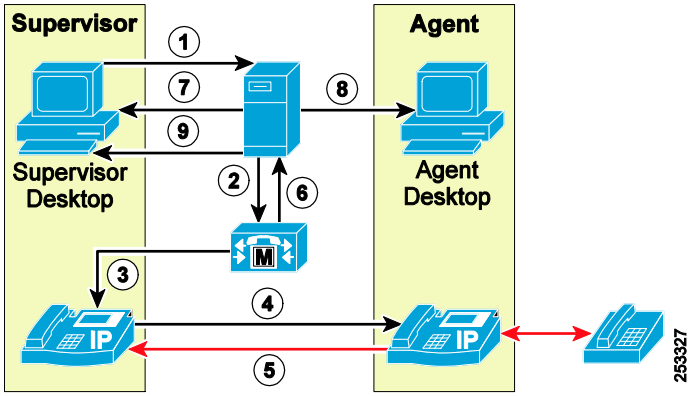
Unified CM Silent Monitor

This section describes how CTI OS accomplishes silent monitoring when the CTI OS Server is configured to use the Unified CM silent monitor.

Unified CCE supports the silent monitoring functionality available in Unified CM 6.0 and higher. Figure 4-6 illustrates the following message flow, which occurs when the Unified CM silent monitor is initiated by the supervisor desktop:

1. The supervisor initiates silent monitoring by sending the Agent.SuperviseCall() message to Unified CCE.
2. Unified CCE sends the Call.startMonitor() message to Unified CM.
3. Unified CM instructs the supervisor's phone to call the built-in-bridge in the agent's phone.
4. The supervisor's phone places the call to the built-in-bridge in the agent's phone.
5. The agent's phone forwards a mix of the agent's and customer's voice streams.
6. Call events for the silently monitored call are sent from Unified CM to Unified CCE.
7. CTI OS sends a SilentMonitorStarted event to the supervisor desktop.
8. CTI OS sends a SilentMonitorStarted event to the agent desktop.
9. CTI OS sends call events for the silently monitored call to the supervisor desktop.

Figure 4-6 Unified CM Silent Monitoring for Unified CCE



Unified CM silent monitoring works the same as other call control functionality provided by Unified CM (such as conference, transfer, and so forth). When Unified CM is used for silent monitoring, a message is sent from the desktop, through Unified CCE, through Unified CM, and out to the phones where silent monitoring is executed.

The messaging through Unified CCE and Unified CM impacts Unified CCE performance. For further details regarding the impact of Unified CM silent monitoring on Unified CCE sizing, see the chapter on Sizing Unified CCE Components and Servers, page 10-1.

Unified CM silent monitoring is supported only for agents who are connected to Unified CCE on the LAN; it does not support mobile agents and remote agents (agents connected to Unified CCE across a WAN).

CTI OS Silent Monitor

This section describes how CTI OS accomplishes silent monitoring when the CTI OS Server is configured to use the CTI OS silent monitor.

The silent monitoring solution provided by CTI Toolkit in Release 7.0 and earlier was integrated in the Client Interface Library (CIL). The CIL had components to capture and forward voice packets as well as components to play back a stream of forwarded voice packets to the supervisor's sound card. This feature limited silent monitoring support to Unified CCE agent desktops deployed behind a Cisco IP Phone and Unified CCE supervisor desktops deployed on the supervisor's desktop.

Starting CTI OS release 7.1, two deployment types are supported: Citrix and Mobile Agent. In these two deployments, the CIL is not deployed where it has access to the voice stream. In Citrix, the CIL is located on the Citrix Server. Agents and supervisors use a Citrix client to run the desktop. When this is done, the desktop runs on the Citrix server. The Citrix client merely displays the UI of the desktop. Because it is the agent's Citrix client that is deployed behind the IP phone, the CIL no longer has access to the voice path. Similarly, it is the supervisor's Citrix client that has the sound card. In this case, the CIL is running on the Citrix server and does not have access to the sound card.

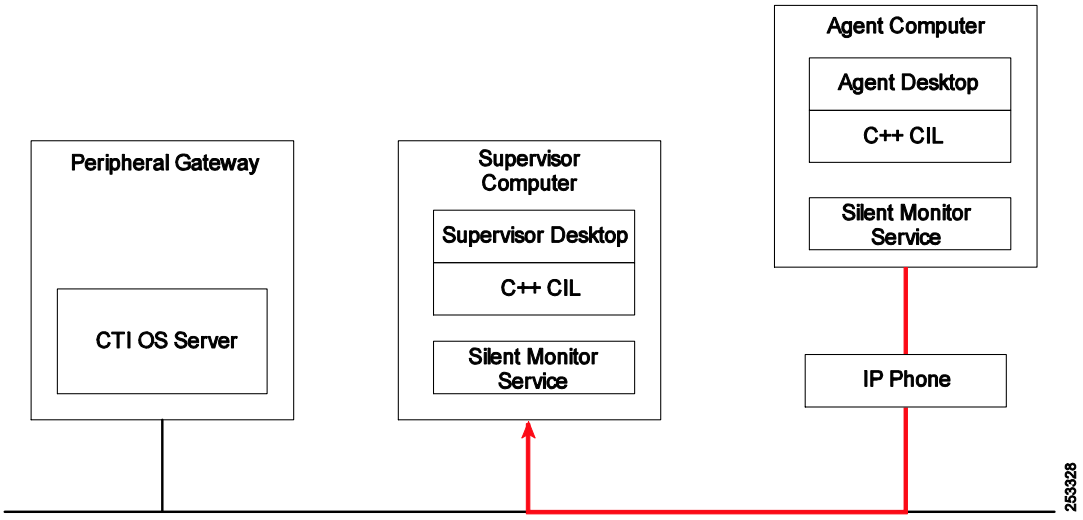
In Mobile Agent deployments, the CIL is deployed on an agent's remote PC. When the agent uses an analog phone, the CIL does not have access to the voice stream.

To support these two deployment models, it was necessary to remove the silent monitor components from the CIL and put them on a separate service. This allows the service to be deployed where it has access to the agent's voice stream or the supervisor's sound card.

The following figures show where the silent monitoring service should be deployed for each deployment model. The red line in each diagram illustrates the path of the monitored voice stream.

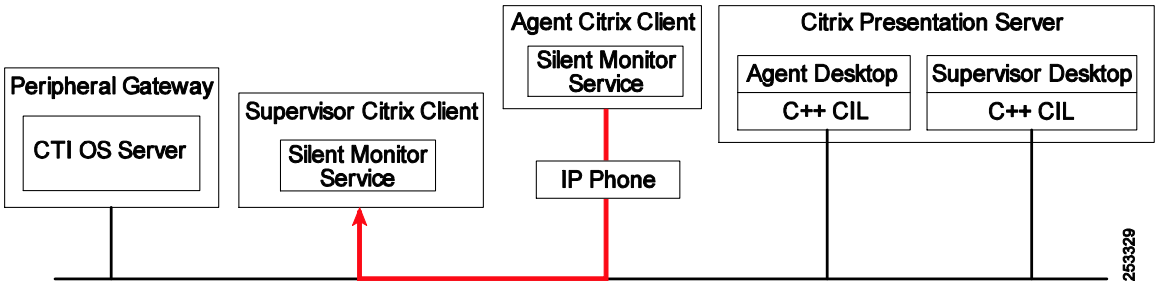
Figure 4-7 and Figure 4-8 illustrate deployments where the agent uses an IP phone. In these deployments, silent monitoring is configured the same way regardless of whether the agent is mobile or not.

Figure 4-7 Silent Monitoring for Cisco Unified CCE When a Mobile or Local Agent Uses an IP Phone



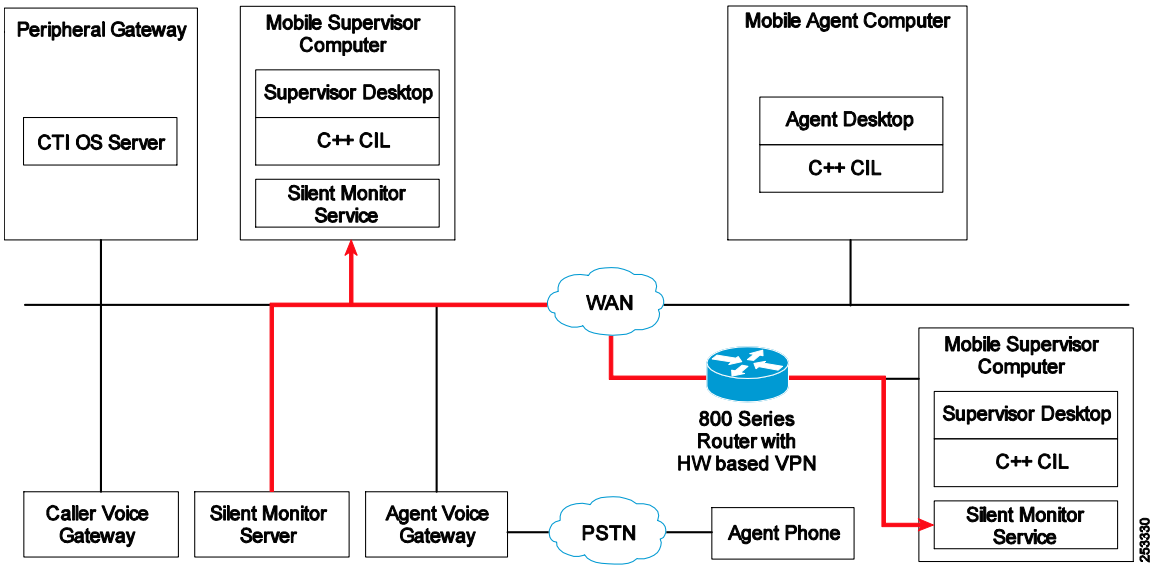
The deployment in Figure 4-7 is very similar to CTI OS Release 7.0 and earlier deployments. The only difference is that the silent monitoring service is running alongside the CIL to provide silent monitoring functionality.

Figure 4-8 Silent Monitoring for Cisco Unified CCE with Citrix When a Mobile or Local Agent Uses an IP Phone



In the deployment model in Figure 4-8, the silent monitoring service is deployed on Citrix clients, where it has access to the agent's voice stream and the supervisor's sound card. The CIL makes a connection to the silent monitoring service and sends it instructions over a TCP connection in order to start and stop silent monitoring sessions.

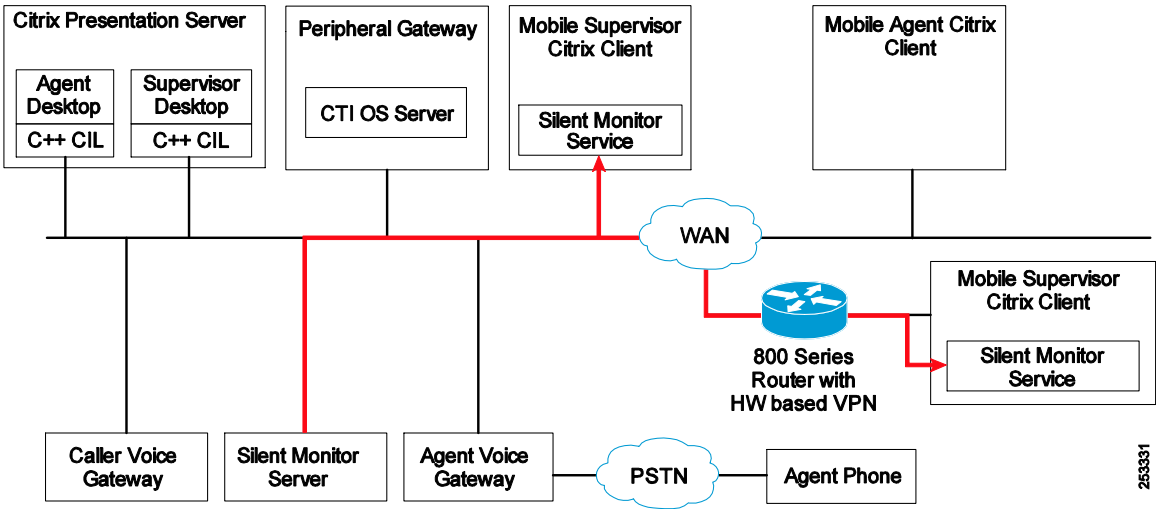
Figure 4-9 Silent Monitoring for a Mobile Agent Using a PSTN Phone



In the deployment model in Figure 4-9, one silent monitoring service is deployed on a switch's SPAN port in order to gain access to voice traffic passing through the agent gateway. This silent monitoring service is used by agents to forward their voice streams to supervisor silent monitoring services.

Supervisors running locally are deployed the same as Unified CCE supervisors. Supervisors running remotely are also deployed the same as Unified CCE supervisors, but a Cisco 800 Series Router with hardware-based VPN is required in order for the supervisor to receive agent voice streams.

Figure 4-10 Silent Monitoring for a Mobile Agent Using a PSTN Phone with Citrix or Microsoft Terminal Services



In the deployment model in Figure 4-10, one silent monitoring service is deployed on a switch's SPAN port in order to gain access to voice traffic passing through the agent gateway. This silent monitoring service is used by agents to forward their voice streams to supervisor silent monitoring services. Mobile agents need to run only their Citrix clients. Agent desktops running on the Citrix server will connect to the silent monitoring server.

Supervisors running locally are deployed the same as Citrix Unified CCE supervisors. Supervisors running remotely are also deployed the same as Citrix Unified CCE supervisors, but a Cisco 800 Series Router with hardware-based VPN is required in order for the supervisor to receive agent voice streams.

In the two mobile agent deployments above (Figure 4-9 and Figure 4-10), calls whose voice traffic does not leave the agent gateway cannot be silently monitored. This includes agent-to-agent calls as well as agent consultations with other agents. The only calls that can be reliably monitored in this case are calls between agents and customers. This is because the mobile agent solution requires separate gateways for callers and agents to ensure that voice traffic is put on the network.

Clusters

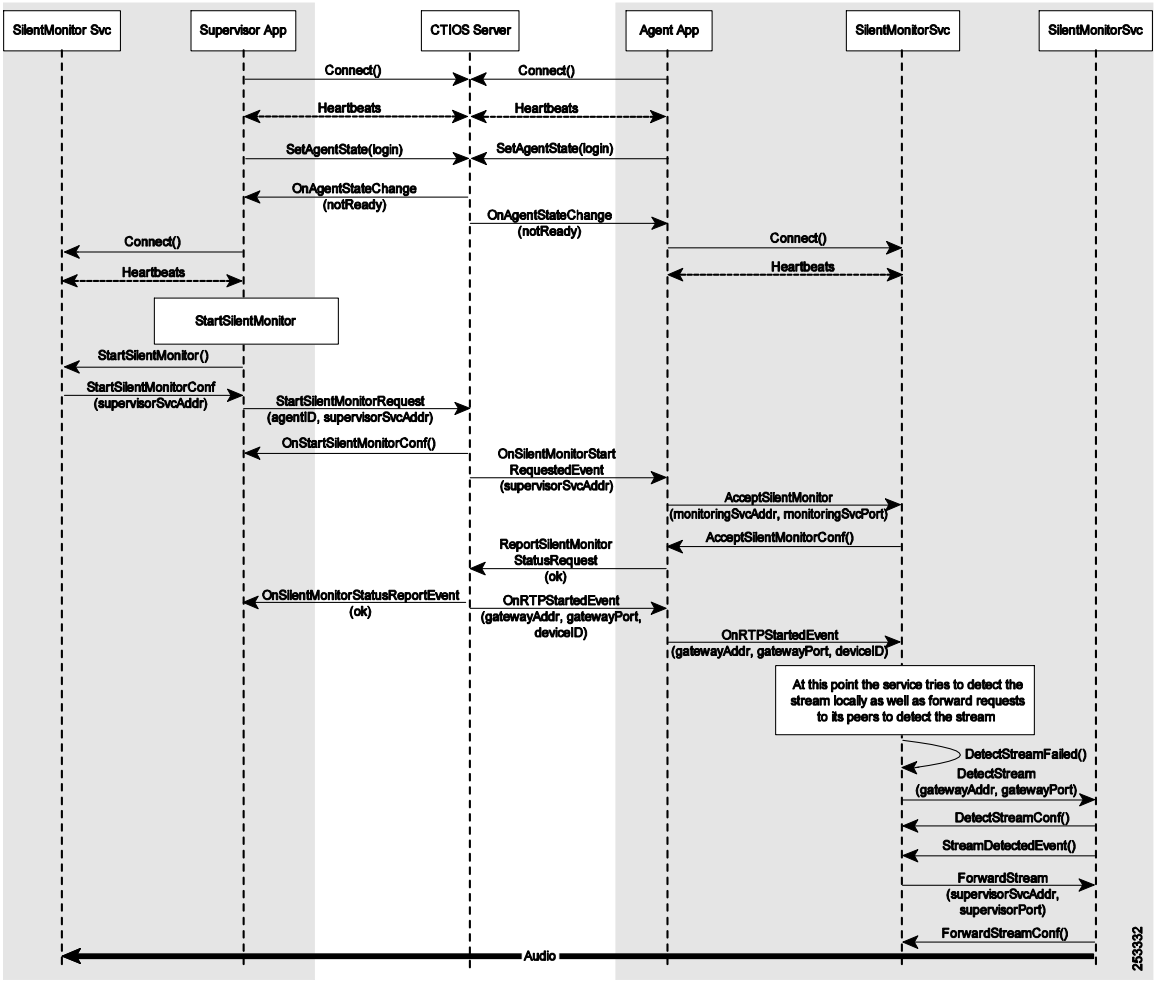
If a mobile agent's login can be handled by one of two gateways, it is possible to cluster two and only two silent monitoring servers together to provide silent monitoring functionality regardless of the gateway that handles the call. A maximum of two silent monitoring servers are supported in a cluster (SPAN) based deployment. When a request to silently monitor the agent is received, the silent monitoring server that receives the request from the agent desktop will forward the request to its peer, and then both silent monitoring servers will attempt to detect the stream. Once the agent's voice stream is detected, it is forwarded to the supervisor's silent monitoring service by the silent monitoring server that detected the stream.

For more information regarding deployment and configuration of the silent monitoring service, refer to the *CTI OS System Manager's Guide*, available on http://www.cisco.com.

Message Flow

Figure 4-11 illustrates the messaging that occurs between the desktops, CIT OS Server, and silent monitoring services when a silent monitor session is initiated. Note that messaging between the desktops and the CTI OS Server has not changed from CTI OS Release 7.0.

Figure 4-11 Message Flow Between Desktops, CTI OS Server, and Silent Monitoring Service



Connection Profiles

In mobile agent deployments, agent desktops learn where and how to connect to their silent monitoring server using a CTI OS connection profile. When an agent logs in, the agent desktop uses the following algorithm to determine where the silent monitoring service is located:

1. If a silent monitoring service is present in the connection profile, attempt to connect to it.
2. If no silent monitoring service is present, determine if the desktop is running under Citrix.
3. If the desktop is running under Citrix, connect to the silent monitoring service running at the Citrix client's IP address.
4. If the desktop is not running under Citrix, connect to the silent monitoring service running at **localhost**.

Supervisor desktops use the following algorithm to find their silent monitoring service:

1. If the desktop is running under Citrix, connect to the silent monitoring service running at the Citrix client's IP address.
2. If the desktop is not running under Citrix, connect to the silent monitoring service running at **localhost**.

If the IPCCSilentMonitorEnabled key is set to 0 in the connection profile, no attempt is made to connect to a silent monitoring service.

CAD Silent Monitoring and Recording

The section describes Cisco Agent Desktop (CAD) silent monitoring.

CAD-Based Monitoring

CAD-based monitoring consists of three types of monitoring:

- Desktop Monitoring
- Server Monitoring
- Mobile Agent Monitoring

Desktop Monitoring

Desktop monitoring uses software running on the agent's desktop (Cisco Agent Desktop) to sniff the network traffic going to and from the agent's phone (hardware phone or software phone) for RTP packets. The monitoring software then sends the RTP packets to the appropriate software over the network for decoding. Desktop monitoring relies on the ability for certain Cisco IP Phones to be daisy-chained with the agent's PC via a network connection and for the phone to send all its network traffic along this connection to the software running on the PC. In this case, the packet sniffing software is able to see the voice traffic coming to and leaving from the agent's phone. It will copy this traffic and send it to the supervisor monitoring the agent or to a recording service for the call to be stored and to be listened to at some later time. Desktop monitoring is not a true service, at least from the perspective of the Service Control Manager. It is a Dynamic-Link Library (DLL), an executable module that is part of Cisco Agent Desktop.

Server Monitoring

Server monitoring uses one or more Cisco Desktop VoIP Monitor Services to sniff the network running over a Cisco Catalyst switch for voice streams. The Cisco Desktop VoIP Monitor Service looks for particular streams to and from phones being monitored or recorded. It then sends the voice packets to the supervisor desktop that is performing the monitoring or to a recording service for storage.

The Cisco Desktop VoIP Monitor Service uses the Switched Port Analyzer (SPAN) or Remote SPAN (RSPAN) monitoring feature of certain Cisco Catalyst switches to sniff the network. The switch uses the monitoring feature to copy the network traffic from one or more sources to a destination port. Sources can be ports and/or Virtual LANs (VLANs). RSPAN allows the source ports to reside on remote switches. The Cisco VoIP Monitor Service connects to the switch via the destination port. This allows the Cisco VoIP Monitor Service to see the voice traffic going to and coming from IP phones.

Mobile Agent Monitoring

Cisco Agent Desktop has the ability to monitor and record mobile agents' RTP sessions by deploying a Cisco VoIP Monitor Service that can see traffic coming from Agent Voice Gateways (this also uses the SPAN feature).

For more information, see the Cisco Agent Desktop product documentation available on http://www.cisco.com.

Fault Tolerance for CAD-Based Monitoring and Recording

Desktop Monitoring

Desktop monitoring is fault tolerant by design. If an agent's desktop fails, only that agent will be unavailable for monitoring and recording.

Server Monitoring and Mobile Agent Monitoring

Server monitoring and mobile agent monitoring are not fault tolerant. If a Cisco Desktop VoIP Monitor Service fails, all agent phones and mobile agent voice gateways associated with that service will be unavailable for monitoring and recording. No backup service can be specified. Monitoring and recording will continue to be available for devices associated with other Cisco Desktop VoIP Monitor Services.

Recording

Recording is fault tolerant. If a recording service fails in a high-availability deployment, the other recording service will assume all recording responsibilities.

Recording Playback

Playback of recordings is not fault tolerant. Recordings are tied to the recording service that captured the recording. If a recording service fails, all recordings associated with that service will be unavailable until it is restored.

Load Balancing for CAD-Based Monitoring and Recording

Desktop Monitoring

Desktop monitoring is load-balanced by design. Monitoring load is distributed between the agent desktops.

Server Monitoring and Mobile Agent Monitoring

Load balancing can be achieved when configuring SPAN ports for, and associating devices with, the Cisco Desktop VoIP Monitor Services. To achieve load balancing, have each VoIP Monitor Service monitor an equal number of agent phones.

Recording

Recording services are selected in round-robin fashion at runtime by the desktops. However, no attempt is made to ensure that the load is balanced between the recording services.

Cisco Remote Silent Monitoring

This section covers Cisco Remote Silent Monitoring. Remote Silent Monitoring (RSM), which allows for the real-time monitoring of agents as a dial-in service.

The RSM solution consists of three components:

- VLEngine
- PhoneSim
- Callflow Script(s) for Unified CVP and IP IVR

For a further description of these components, refer to the *Cisco Remote Silent Monitoring Installation and Administration Guide*, available at http://www.cisco.com.

Hardware Considerations

The RSM solution is highly integrated part of a Cisco Unified Contact Center Enterprise environment. Because of this, the functioning of RSM requires resources from various other components of the platform as a whole. To properly integrate RSM, then, requires an understanding of its interactions with the rest of the environment so that capacity can be properly planned, provisioned, and managed.

Platform Considerations

In particular, RSM interacts mainly with the Unified CM cluster.

The RSM server has two tie-ins with each Unified CM cluster in the environment that it is configured to use:

Simulated Phones: RSM's PhoneSim component requires that a Cisco Unified IP Phone 7941 device entry be created on the Unified CM cluster for each of the simulated phones (or "simphones") it is configured to manage. For instance, a RSM system that is configured to handle up to 100 dialed-in supervisors monitoring agents on a particular Unified CM cluster will need to have at least of these 100 simphones. To the Unified CM cluster itself, these simphones appear as normal Cisco Unified IP Phone 7941 SIP phones; however, in reality they are homed to and controlled by PhoneSim instead of being an actual physical phone device.

When compared with the usage profile of a normal phone, the simphone usually puts a lighter load on the Unified CM cluster. This is because it exhibits only a small set of behaviors, consisting of:

- Registering with the Unified CM cluster when PhoneSim is started.
- Making a "monitoring call" to an agent's phone when a dialed-in supervisor requests to monitor that agent. The agent's phone then forks off a copy of the conversation the agent is having to the simphone.

JTAPI: When RSM is integrated into the environment, a JTAPI user is created and associated with each agent phone device that can be monitored, as well as with each simphone device that was created on the cluster.

When an agent is to be monitored, a JTAPI monitor request call is made from the RSM server to the Unified CM cluster that manages that agent's phone. Also, while RSM is in use, a JTAPI CallObserver is kept attached to each simphone device. It is also attached to an agent phone device, but only while the JTAPI monitor request is being issued to that device.

JTAPI connections may optionally be encrypted. However, this will induce a slight performance penalty on the server itself when higher agent loads are utilized. For more information on enabling JTAPI connection security, refer to the *Cisco Remote Silent Monitoring Installation and Administration Guide*, available at http://www.cisco.com.

CTI OS Server

RSM makes a persistent "monitor-mode" connection to each CTI OS server it is configured to use. Through this connection certain platform events such as call start, call end, agent on hold, and so forth, are streamed in real-time.

Besides this, RSM will make an additional, short-lived "agent-mode" connection to possibly each CTI OS server when a supervisor dials in and authenticates. The purpose of this connection is to validate the supervisor's entered credentials by performing a corresponding login into CTI OS. Note that, if the built-in authentication mechanisms of the RSM callflow (for example, the checkCredentials API call) are not used, this connection is not made. If the login is successful, that supervisor's team membership is requested by the RSM server. Once returned, a logout is called and the connection is terminated.

Note that the total supervisor count in Unified CCE must be spread across CTI OS desktop users and RSM. For example, in a 2000 agent configuration, up to 200 agents can be supervisors. This means that the total supervisor count between CTI OS and RSM must not exceed 200.

CTI OS connections may be optionally encrypted (via use of IP Sec configurations). However, this will induce a significant performance penalty on the server itself when higher agent loads are utilized. For more information on enabling CTI OS connection security, refer to the *Cisco Remote Silent Monitoring Installation and Administration Guide*, available at http://www.cisco.com.

VRU

The RSM platform does not directly media-terminate inbound calls. Instead, supervisors dial into a Unified CVP or IP IVR-based VRU system, which runs call flow script logic that interacts with services hosted on the RSM server via HTTP. Thus, if a given RSM installation is to support up to 40 dialed-in supervisors, there must be a VRU present (as well as the necessary PRI/network resources) that can offer this same level of support.

Furthermore, a caller accessing RSM will often place a higher load on the VRU's processor(s) and memory than a caller accessing some more traditional IVR-type callflow. This is because, in a more traditional IVR callflow, shorter, oftentimes cached or non-streamed prompts are played, separated by periods of caller input gathering and silence. With RSM, however, the predominant caller activity is monitoring an agent's call, and to the VRU this looks like the playback of a long streaming audio prompt, which is an activity that requires a relatively high level of VRU processor involvement.

With Unified CVP deployments, supported VXML gateway models are listed in the *Hardware and System Software Specification for Cisco Unified Customer Voice Portal (Unified CVP)*, otherwise known as Unified CVP Bill of Materials (BOM), available at http://www.cisco.com.

When provisioning a VRU for use by RSM, a good rule of thumb is to count each RSM call as 1.3 non-RSM calls on a processor/memory-usage basis. For example, for a VRU that can normally handle 40 concurrent calls, plan for it to be able to handle only 30 RSM calls. ((40 / 1.3) = 30)

Also note that RSM makes extensive use of VXML Voice Browser functionality under both Unified CVP and IP IVR.

When RSM is used with CVP, the gateway 'IVR prompt streaming for HTTP' needs to be enabled. Note that this setting is not recommended for other CVP applications. Therefore RSM requires a dedicated VXML gateway. This gateway must not be used for other CVP applications. Also, 1GB of gateway memory is recommended for use with RSM. This will support up to 40 concurrent monitoring sessions per gateway.

Agent Phones

Use of RSM to monitor an agent requires that that agent's phone be a third generation of Cisco Unified IP Phones 79*x*1, 79*x*2, 79*x*5, 7970, or newer. This is because these phones include extra DSP resources in the form of a Built-in-Bridge (BiB). The BiB allows the phone to fork off a copy of the current conversation stream to the RSM server.

Cisco Unified Contact Manager provides for a maximum of one active monitoring session per agent because the agent's phone can handle only one active monitoring session and one active recording session at any given time.

So, if a third-party recorder is recording the agent's conversations, the agent can still be monitored by a supervisor using supervisor desktop or RSM. However, if both a RSM-based supervisor and a supervisor-desktop-based supervisor both tried to monitor the agent during the same time period, the request would fail with the last one to try because it would exceed the above-mentioned monitoring limit.

Note that RSM will set up only one monitoring session through Unified CM for a single monitored agent, even if two or more RSM users are requesting to monitor the agent's call at the same time. In this case, RSM forks the stream to cover all RSM users. This allows more than two RSM-based supervisors to monitor the same agent, for instance. However, if there are multiple RSM servers in the environment that monitor the same agent, they will each make a separate monitoring call to that agent.

If the monitoring call limit has been reached for a specific agent and a dialed-in supervisor then attempts to monitor this same agent, the supervisor's request will be denied via an audio prompt feedback from the system stating that the agent cannot be monitored.

RSM Hardware Considerations

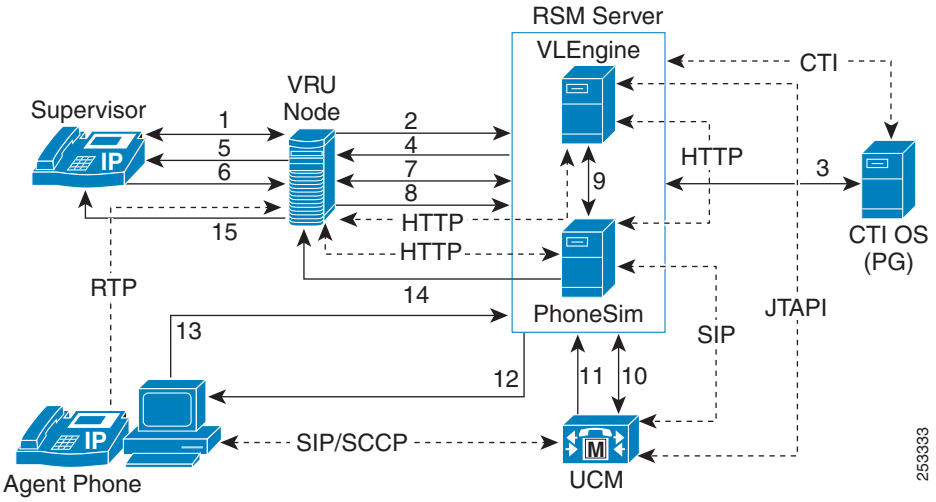
RSM is supported in installations where the number of agents in the enterprise is less than 8,000 and the number of maximum concurrent number supervisors using the system is less than 80. In all supported RSM configurations, the VLEngine and PhoneSim components are installed on the same physical server.

For more information, refer to the RSM Requirements section of the *Cisco Remote Silent Monitoring Installation and Administration Guide*, available at http://www.cisco.com.

RSM Component Interaction

Figure 4-12 illustrates the types of interactions that occur when a supervisor dials into an RSM-enabled platform and monitors an agent.

Figure 4-12 Remote Silent Monitor Enabled Call Flow



RSM Call Flow

Figure 4-12 shows the following call flow steps:

1. Supervisor calls in, and the call is media-terminated on the VRU (Unified CVP or IP IVR). The VRU runs the RSM callflow script to handle the call. The call begins by the user being asked to authenticate himself or herself. The user then enters his or her credentials.
2. After the user enters his or her credentials, the VRU makes a login request to RSM over HTTP.
3. The VLEngine component in RSM interacts with the CTI OS server to validate the authentication credentials.

4. VLEngine replies back to the VRU node via HTTP with the authentication result.
5. If the supervisor is successfully authenticated, the script in the VRU will play the main menu prompt. From here, the supervisor will be allowed to monitor an agent.
6. The supervisor chooses to monitor a single agent from the main menu, and enters a Directory Number (DN) of an agent to be monitored.
7. The VRU checks with VLEngine if the given agent can be monitored. VLEngine then checks whether the agent with that DN is logged in, is in talking state, and is in the supervisor's team, using previously cached event feed information from the CTI OS server. If so, it replies back to the VRU node.
8. The VRU node then sends a monitor request to PhoneSim to monitor the entered DN.
9. VLEngine works internally using HTTP.
10. Following that, VLEngine sends a JTAPI request to Unified CM to monitor the agent's phone, and it gets a JTAPI success response.
11. The PhoneSim component will then receive a SIP-based instruction from Unified CM for a simulated phone that it manages, to establish a monitoring call with the agent's phone.
12. The chosen simulated phone establishes the monitoring call with the agent's phone based on Unified CM's above request.
13. After the establishment of a monitoring call from RSM server to agent, the agent phone's Built-in-Bridge (BiB) forwards the call conversation to PhoneSim in the form of RTP packets.
14. In turn, PhoneSim strips the RTP headers and streams this data to the VRU node over HTTP as a response to the request made earlier in step 8.
15. The VRU then plays the data to the supervisor as if it were a streaming audio prompt.

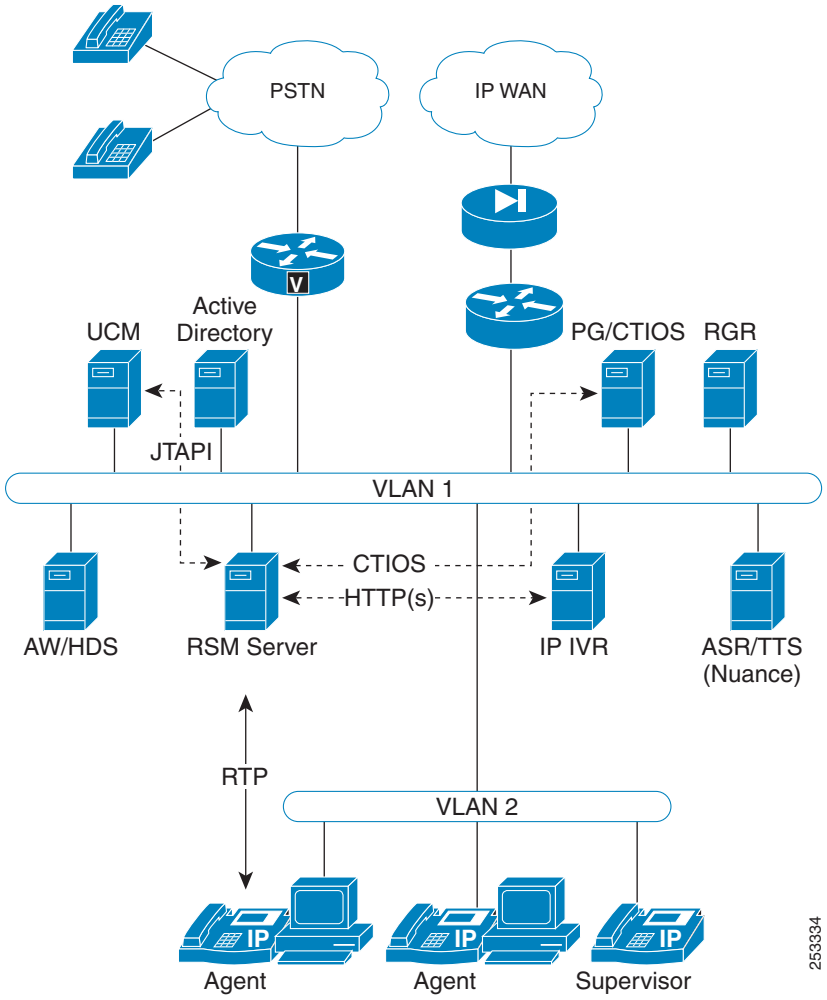
Deployment Models

This section illustrates some basic supported RSM deployments.

Single Site

Figure 4-13 illustrates the basic network connectivity of RSM deployed within a typical single-site configuration.

Figure 4-13 Typical RSM VLAN Configuration



As shown in Figure 4-13, supervisors may dial in through a VoIP phone as well as through the PSTN. The VRU that handles the supervisor's call is IP IVR in this case.

Figure 4-13 also illustrates the various protocol interfaces that RSM has into the rest of the system:

- **HTTP(S)**: As stated previously, HTTP is used as the carrier protocol for VRU-based requests into the RSM system. A request takes standard URL form and may look like one of the following URLs:

  http://rsmserver:8080/vlengine/checkUserCredentials?supervisorID=1101&pin=1234&outputFormat=plain

  http://rsmserver:8080/vlengine/canMonitorAgentID?supervisorID=1101&agentID=1001&outputFormat=vxml

  The first request about is for the checkUserCredentials API call, while the second is for the canMonitorAgentID API call. Parameters to these requests are passed via the GET method. The return data (as an HTTP response) is either plaintext or encapsulated in VoiceXML, depending on the API call being used and on the value specified for the outputFormat parameter (if available for that call).

- **CTI OS**: The RSM server makes several connections to CTI OS. One of these connections is for receiving platform events. (In the language of CTI OS, it is a monitor-mode connection.) The other(s) are what CTI OS calls agent mode connections and they are used to authenticate logging-in supervisors if the standard authentication facilitates are being utilized.

- **JTAPI**: The request to start monitoring an agent's phone is made through JTAPI. This requires a JTAPI application user to be defined on each Unified CM cluster in the environment, and to be associated to all agent phones.
- **RTP**: While a dialed-in supervisor is monitoring an agent, there will be a monitoring call in progress from the BiB (built-in-bridge) of that agent's phone to the RSM server. While the signaling data for this call is run through Unified CM (just like any other call), the RTP traffic will flow between the agent phone and the RSM server.

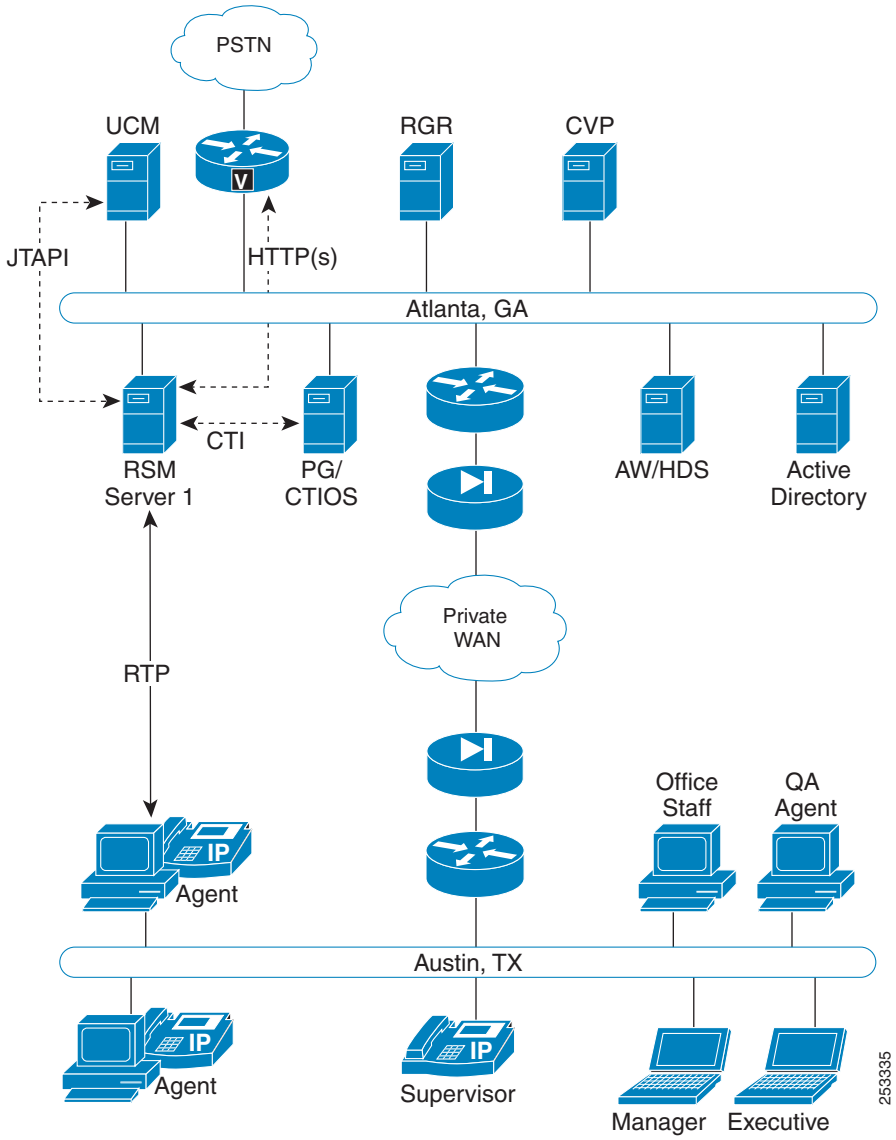
Multisite WAN

The follow scenarios depict basic supported configurations for the RSM product in a multi-site deployment.

Single Cluster, Single VRU

Figure 4-14 depicts a simple multi-site setup involving a single Unified CM cluster and a single VRU.

Figure 4-14 Multi-Site Deployment with a Single Unified CM Cluster and Single VRU



In this case, the Unified CM and Unified CCE environment is co-located in Atlanta, and the Austin location contains the entire end-user population. The VRU is a VXML Gateway/voice gateway in Atlanta, controlled by a Unified CVP Call Server also in Atlanta.

The supervisor in Austin could possibly have two ways of dialing into the RSM system:

- Through the PSTN — Here the supervisor would dial an E.164 number, and the call would be hairpinned through the voice gateway. The Unified CVP RSM callflow application would handle the call as normal from that point.
- As a VoIP extension — In this case, Unified CM would have a trunk configuration set up to the VRU. The call would remain VoIP all the way through, and the call would likewise be handled by the Unified CVP RSM callflow application.

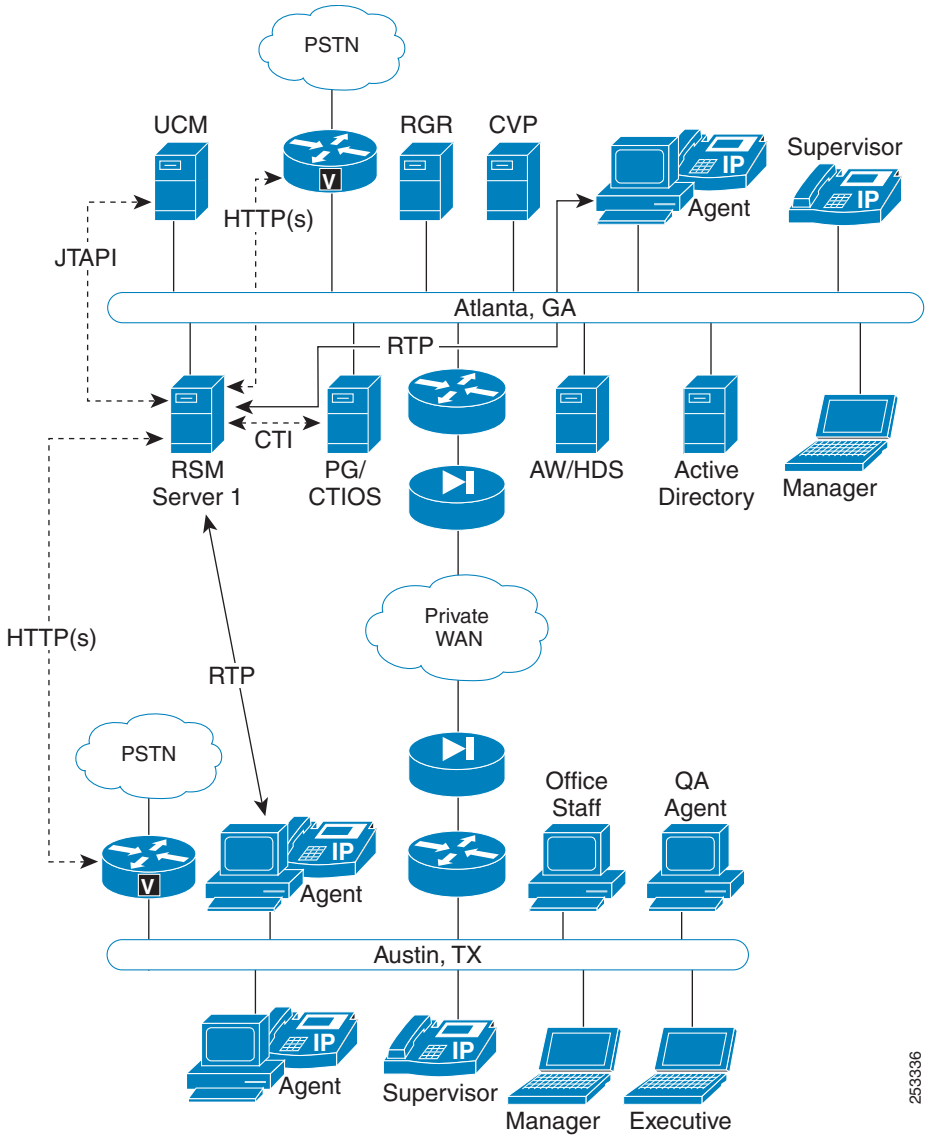
In this scenario, all RSM traffic is confined to the Atlanta site except:

- The RTP traffic of the agent being monitored (signified as a red dotted line)
- The actual supervisor call into the platform

Single Cluster, Multiple VRUs

Figure 4-15 depicts a multi-site deployment with a single Unified CM cluster and multiple VRUs.

Figure 4-15 Multi-Site Deployment with a Single Unified CM Cluster and Multiple VRUs



This scenario is similar to the previous one, with the addition of PSTN access at the Austin site. This scenario also adds personnel to the Atlanta site.

With the addition of a PSTN egress point in Austin, a call from a supervisor at the Austin location to the RSM system could be backhauled across the WAN (if VoIP end-to-end) or sent across the PSTN if the Atlanta DID associated with the RSM application was dialed.

In this example, Unified CVP is still used as well as the Unified CVP Call Server. However, there are two VXML Gateways, one at each site. The environment is configured so that a supervisor dialing RSM from Austin will be routed to the RSM callflow application on the Austin VXML Gateway, while a supervisor dialing in from Atlanta will be routed to the Atlanta VXML Gateway.